

MULTIPLE TIME SCALE APPROACH
TO HIERARCHICAL AGGREGATION OF
LINEAR SYSTEMS AND FINITE STATE MARKOV PROCESSES

by

Marcel Coderch i Collell

S.B., Universitat Politècnica de Barcelona, (SPAIN)
(1976)

S.M., Massachusetts Institute of Technology
(1979)

E.E., Massachusetts Institute of Technology
(1979)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

July 1982

© Marcel Coderch, 1982

Signature of the Author
Department of Electrical Engineering and Computer
Science, July 20, 1982

Certified by
Alan S. Willsky
Thesis Supervisor

Accepted by
Arthur C. Smith
Chairman, Departmental Graduate Committee

The author hereby grants to M.I.T. permission to reproduce and to distribute copies of this thesis document in whole or in part.

MULTIPLE TIME SCALE APPROACH TO HIERARCHICAL
AGGREGATION OF LINEAR SYSTEMS AND FINITE
STATE MARKOV PROCESSES

by

Marcel Coderch i Collell

Submitted to the Department of Electrical Engineering and
Computer Science on July 20, 1982 in partial fulfillment
of the requirements for the
Degree of Doctor of Philosophy

ABSTRACT

The use of weak couplings and time scale separation as tools for model simplification is considered. Unifying and conceptually clarifying existing approaches, a new problem formulation is proposed and a uniform asymptotic approximation to $\exp\{A(\epsilon)t\}$ (under a certain multiple semistability condition) is developed which clarifies the relationships among weak couplings, singular perturbations, time-scale separation, aggregate models and asymptotic approximations.

As a result of this asymptotic approximation, an algorithm for the calculation of a hierarchy of aggregated models of a given linear system is obtained.

A more detailed analysis is conducted for two special class of models: finite-state Markov processes (FSMP's) with rare events, and linearized "swing" equations for electric power networks. It is shown that any singularly perturbed FSMP can be hierarchically aggregated and that in turn, the aggregated models can be combined to produce an approximation of the original process uniformly valid over $t \in [0, \infty)$. The properties of stochastically discontinuous FSMP's are analyzed and interpreted as the limiting behavior of singularly perturbed FSMP's. Coherence phenomena in electric power networks are explained in terms of weak interactions among groups of generators and the determination of coherence areas is shown to be equivalent

to the aggregation of FSMP's. Coherence-area based models are shown to be asymptotic approximations of a peculiar kind.

Finally, the use of aggregated models in filtering of singularly perturbed FSMP's is explored together with the trade-offs involved in the detection of rare events.

Thesis Supervisor: Alan S. Willsky

Title: Associate Professor of Electrical Engineering and
Computer Science

ACKNOWLEDGEMENT

I am grateful to Professor Alan S. Willsky for this enthusiasm, guidance and encouragement during the course of this research. His insights, observations and suggestions have contributed decisively to the success of the work reported here.

I thank my thesis readers, Professor Shankar S. Sastry, Professor Steven I. Marcus and Dr. David Castañon, for their valuable guidance and suggestions during the course of this research.

Professor Sanjoy K. Mitter with his uncompromising commitment to scientific inquiry has been a source of inspiration for me. I thank him for that.

My stay at MIT would have not been so fruitful had not been for the thought-provoking, marvelous lectures and conversations with Noam Chomsky, David Noble and Louis Kampf. They helped me develop an awareness of the social dimensions of our work and of the role of intellectuals and engineers in society. This has been of outmost value to me.

I wish to express my gratitude to Dr. Howard Chizeck for his friendship and for many valuable conversations. Thanks also go to my other fellow students, Yehuda Avniel, Rickly Lee, Jim Lewis, Charlie Rohrs and Peter Thompson. I benefited in many ways from hours of active and stimulating discussion with them.

Thanks to Fifa Monserrate for her excellent typing of this thesis and for her friendship, and to Arthur Giordani for the drafting.

I gratefully acknowledge the continuing support of the Fundación ITP (SPAIN) that made possible my stay at MIT.

Finally, to my wife Isabel, special gratitude. For her companion, affection and love. For being there when I needed most. For taking more than her share in raising our daughters, Belona and Anna. Without her this thesis would have never been completed.

"The so-called concern about Science and Technology *per se* -the belief that they are value free and politically neutral, and that their "advancement" is a good and desirable thing because knowledge *can* always be put to good uses, even if it is not, presumably- is nothing but an *ideology* of self-justification which tries to hide the subservience of science and technology -in their priorities, their language, and their utilization- to the demands of capitalist institutions and domination.

Technical and scientific culture and competence bear the mark of a social division of labor which denies to all workers, including the intellectual ones, the insight into the system's functioning and overall purposes, so as to keep decision-making divorced from productive work, conception divorced from execution, and responsibility for producing knowledge divorced from responsibility for the uses knowledge will be put to."

Andr  Gorz

TABLE OF CONTENTS

	<u>PAGE</u>
CHAPTER I: INTRODUCTION	13
1.1 Motivation and Goals	13
1.2 Description of Results	21
1.3 Thesis Outline	28
CHAPTER II: PREVIOUS WORK AND RELATED LITERATURE	30
2.1 Introduction	30
2.2 Asymptotic Analysis of Singularly Perturbed Differential Equations	31
2.3 Markov Process Aggregation	38
2.4 Order-Reduction and Aggregation Methods in Filtering and Control	40
CHAPTER III: MATHEMATICAL PRELIMINARIES	44
3.1 Introduction	44
3.2 Linear Operators in Finite Dimensional Spaces	44
3.2.1 Normal Vector Spaces	44
3.2.2 Linear Operators	45
3.2.3 Analysis with Operators	48
3.2.4 The Resolvent	50
3.2.5 Operators with Semisimple Null Structure	54
3.3 Asymptotic Analysis	57
3.3.1 Orders of Magnitude	58
3.3.2 Asymptotic Expansion	60
3.4 Perturbation Theory in Finite Dimensional Spaces	61
3.4.1 The Problem	61
3.4.2 Perturbation of the Eigenvalues	63

	<u>PAGE</u>
3.4.3 Perturbation of the Resolvent	65
3.4.4 Perturbation of the Eigenprojections	67
3.5 Linear Dynamical Systems	72
3.5.1 Stability and Semistability	73
3.5.2 Flow-Invariant Subspaces and Splitting of Evolutions	75
3.5.3 Positive Linear Systems	76
CHAPTER IV: MULTIPLE TIME SCALE BEHAVIOR OF SINGULARLY PERTURBED LTI SYSTEMS	82
4.1 Introduction and Overview	82
4.2 Singularly Perturbed LTI Systems	83
4.2.1 Regular and Singular Perturbations, and Time Scales	83
4.2.2 Eigenvalues and Time Scales	84
4.3 Literature Survey	85
4.4 The Multiple Semistability Case: Complete Time Scale Decomposition	98
4.4.1 The Multiple Semisimple Null Structure and the Multiple Semistability Condition	98
4.4.2 Uniform Asymptotic Approximation of $\exp\{A_0(\epsilon)t\}$	116
4.4.3 Multiple Time Scale Behavior and Reduced-Order Models	120
4.4.4 Necessity of the Multiple Semistability Condition	127
4.4.5 Computation of the Reduced-Order Models	139
4.5 Partial Time Scale Decomposition	149
4.5.1 Uniformly Stable Systems	151
4.5.2 Non-Uniformly Stable Systems	156
4.6 Summary and Conclusions	162

	<u>PAGE</u>
CHAPTER V: HIERARCHICAL AGGREGATION OF FINITE STATE MARKOV PROCESSES	166
5.1 Introduction and Overview	166
5.2 Literature Survey	166
5.3 A Motivating Example	171
5.4 Stochastically Discontinuous FSMP's	177
5.4.1 Implications of Stochastic Discontinuity	183
5.4.2 Aggregation of Stochastically Discontinuous FSMP's	187
5.5 Singularly Perturbed FSMP's and Aggregated Models	194
5.6 Example	203
CHAPTER VI: TIME SCALES AND COHERENCE AREAS IN POWER SYSTEMS	212
6.1 Introduction	212
6.2 Linearized Swing Equations	214
6.3 A Motivating Example	221
6.4 Multiple Time Scale Behavior of Linearized Swing Equations	232
CHAPTER VII: AGGREGATE FILTERING FOR SINGULARLY PERTURBED FSMP's	239
7.1 Introduction	239
7.2 The Optimum Filter Equations: Qualitative Analysis	246
7.3 The Hypotesis Testing Problem	257
7.3.1 Boundary Classification	260
7.3.2 Probability of Error Versus Mean Time to Detection	262
7.3.3 Probability of Large Deviations	264

	<u>PAGE</u>
7.4 Detection of Rare Transitions	266
7.4.1 The Effect of the Perturbation: Boundary Classification	267
7.4.2 Mean Time Between False Alarms versus Mean Time to Detection	273
7.5 Hierarchical Filtering	282
7.5.1 Filtering with Aggregated Measurements	282
7.5.2 Filtering with Aggregated and Decentral- ized Measurements	295
Appendix 7.A One dimensional diffusions in bounded domains: Definitions and Basic Results	303
Appendix 7.B Expected time to detection for the hypotesis testing problem	312
Appendix 7.C Mean exit time from the boundary	314
Appendix 7.D Mean time between false alarms	317
CHAPTER VIII: CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH	319
REFERENCES	328

LIST OF FIGURES

	<u>PAGE</u>
FIGURE 1.1	18
1.1	18
1.2	18
1.3	18
1.4	20
1.5	20
1.6	20
FIGURE 4.1:	110
4.1	140
4.2	164
4.3	164
FIGURE 5.1:	172
5.1	173
5.2	174
5.3	174
5.4	176
5.5	204
5.6	205
5.7	207
5.8	209
5.9	211
FIGURE 6.1:	213
6.1	222
6.2	222
FIGURE 7.1:	241
7.1	241
7.2	248

	<u>PAGES</u>
FIGURE 7.3 Potential for $\rho(t)=1$ and $\rho(t)=0$	253
7.4 Qualitative performance measures	255
7.5 The Hypotesis Testing Problem	258
7.6a The drift coefficient	269
7.6b Potential and noise intensity for the rare transitions case	269
7.7 Scale density	270
7.8 Steady state probability density assuming $\rho^E(t)=1$ for $t \geq 0$.	274
7.9 The process $\rho^E(t)$ considered in Section 7.5.1 and its aggregate approximation $\hat{\rho}^E(t)$	283
7.10 Optimum filter with perfect aggregate observations	297
7.11 Hierarchical filter	298
7.12a True process and (b) erroneous model	301

CHAPTER I: INTRODUCTION

1.1 Motivation and Goals

An important part of today's engineering activity is devoted to the development and understanding of models for the operation of complex interconnected large-scale systems. The process of modelling is always difficult because, as with any abstraction, it involves a compromise between accuracy of representation and complexity. Overly detailed and comprehensive models are generally of little use since they often lead to intractable problems.

The predicament of the system's engineer is to overcome this difficulty by devising relatively simple models and techniques of analysis and design which, in an approximate way, still capture the relevant factors to be considered in a given situation and which lead to a satisfactory design or explanation of the system under consideration. In the course of his work the engineer develops a "feeling" for which are the relevant factors in different problems, and using heuristic, sometimes ill-defined, concepts, manages to arrive at solutions that, at least under normal circumstances, meet the requirements. The articulation and systematization of this tacit knowledge into a "theory" is usually a slow and intricate process, the development of which traditionally has been the role of academia.

Consider for example one of the most complex man-made systems: electric power networks. Hundreds of generators are tied together with hundreds of thousands of conduction lines supplying energy to millions of users who can switch loads at will. The analysis, design and operation of such a system involves many individuals none of whom has a complete picture of the whole system. The engineer designing the control for the turbines of a given generator, for example, views the rest of the system as something which is not affected by the workings of his controller whose mission is to keep vapor pressure and temperature within specified bounds. In contrast, the operator at a central control room is concerned about daily load fluctuations and the distribution of load among different power generating stations. For him the turbine that worried our first engineer no longer exists as it is a microscopic detail in his view of the system. At a different level, we can imagine the maintenance department or the board of directors deciding on inspection requirements or new installations. Daily scheduling is not of detailed concern with regards to these issues because they deal with aspects of the system's behavior which are relevant only over much longer periods of time.

As another example, consider the economic system. One could attempt to deal with individual producers and consumers and trace all interactions among different economic agents. This being an obvious impossibility, however, we rapidly turn to the use of aggregated

variables such as the total output of different economic sectors, aggregate consumption or aggregate investment, and the objective is then to find "macroeconomic laws", i.e., a model in terms of the aggregate variables.

Two basic concepts are routinely used in the modelling process mentioned above: *weak couplings and time scale separation*. A system is heuristically decomposed into subsystems which are assumed, on the basis of practical experience or physical considerations, to interact weakly. That is, changes in one of the subsystems have only minor effects on other parts of the system. This notion leads to the classification of variables into a smaller number of groups. The interactions within each group are studied as though interactions among groups did not exist and the interaction among groups is described by means of aggregate variables without regard for interactions within each group.

Along with this "spatial" aggregation usually goes a notion of time scale separation. Some phenomena are known to be very fast while others manifest themselves only on the long run. Accordingly, in studying the fast phenomena the slow variables are taken as constants while in the analysis of the behavior of the system over long periods of time, only the average effect of the fast variables, and not the details of their evolution, are assumed to be of relevance.

As pointed out in [Sim 61], this reasoning, when explicitly spelled out, seems rather bold and yet we implicitly use it

every time we deal with a large-scale system. To date, there is no body of theory that would explain in a satisfactory way the following questions:

- a) under what conditions is this approach justified?;
- b) what is the nature of the approximation involved?
(clearly the conditions for exact aggregation are very severe and restrictive;
- c) what rules and criteria can we use to determine what variables can be aggregated together?, and
- d) how satisfactory, in terms of error or loss of performance for example, can we expect this approximation to be?

Building upon the work of many researchers, this thesis represents a contribution to the on-going effort directed at filling this void. The fundamental objectives of the research reported here have been:

- a) to formalize the concepts of weak couplings, time scale separation and aggregate modelling;
- b) to determine conditions under which a linear time invariant system is amenable to aggregate analysis;
- c) to develop a systematic methodology for the derivation of the aggregated models; and

- d) to explore how problems posed for large-scale systems can be approximately solved using aggregated models.

To motivate the specific problem formulation used for the purposes outlined above we discuss two simple examples. Consider first the finite state Markov process $\eta^\varepsilon(t)$ shown in Figure 1.1 as an example of a system with a structure that is clearly amenable to the aggregation philosophy we have just described. The process $\eta^\varepsilon(t)$ can be thought of as modelling a system with the following characteristics: it may operate in two different modes corresponding to the two sets of states $X_1 = \{1,2\}$ and $X_2 = \{3,4\}$. Some rare event (notice $\varepsilon \ll \lambda_1$) results in a change in the operating mode, i.e., in a transition between X_1 and X_2 . If the matter of interest is the detailed behavior of the system when in a given mode, and the period of concern is small compared to $1/\varepsilon$, then the model obtained by putting $\varepsilon=0$, as in Figure 1.2, is a satisfactory approximation. If, on the contrary, one is mainly concerned with the long run characterization of changes in the mode of operation, then the process in Figure 1.3 seems to be an adequate simplification, provided the aggregated parameters λ' and λ'' are properly adjusted. As we will see, this reasoning can be rigorously founded and, more importantly, we will show that although the two simplified models in Figures 1.2 and 1.3 give only partial

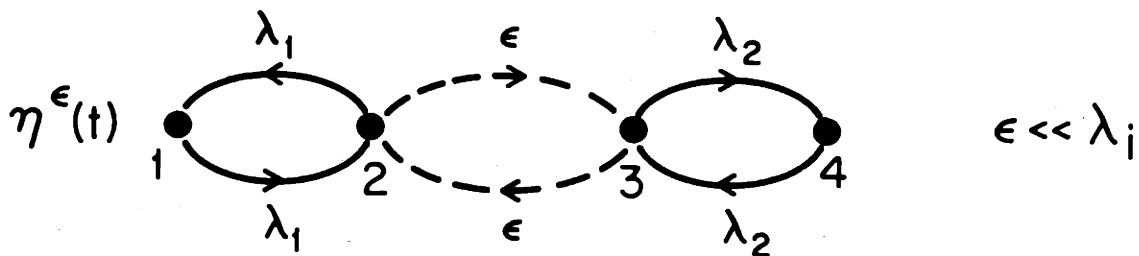


Figure 1.1: A finite state Markov process with diagonally dominant structure.

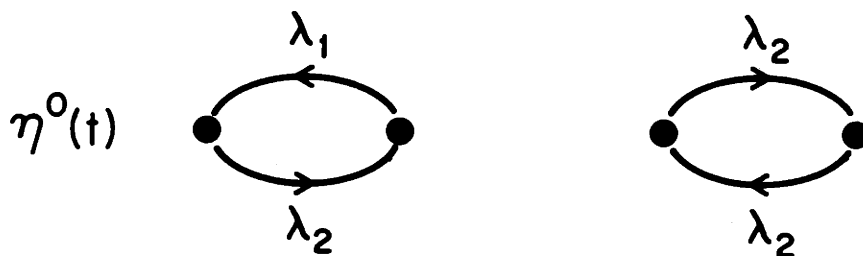


Figure 1.2: The decoupled approximation.

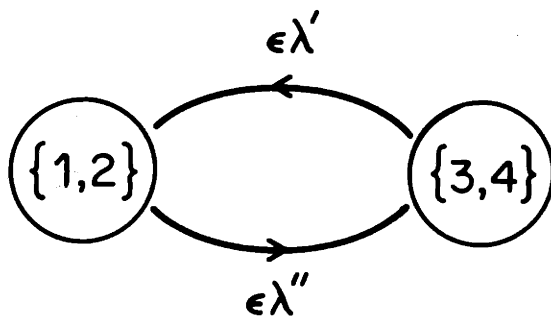


Figure 1.3: Approximation of $\eta^\epsilon(t)$ based on aggregation.

pictures of the behavior of $\eta^\varepsilon(t)$, they can be combined to produce an approximation that captures all aspects of the evolution of $\eta^\varepsilon(t)$.

The example above is a trivial one in a crucial aspect: a quick look at Figure 1.1 reveals how the system should be partitioned and what should be the structure of the aggregated model. In fact, linear systems with "N diagonally dominant blocks" have for some time been the object of study in large scale systems because of the ease with which a meaningful decomposition of the system can be found. This assumption, however, begs all questions. A fundamental goal of this thesis is to do away with this requirement, and to develop a methodology for aggregation based on a deeper understanding of the causes and effects of weak couplings and time scale separation.

Consider next the electric network shown in Figure 1.4 that was proposed as a suggestive example in [Kok 80]. This system has two eigenvalues of order $-R$ and two more of order $-r$ giving an indication of time scale separation: some modes are fast and some are slow (the former approximately decay as e^{-Rt} , the latter as e^{-rt}). In the search for a simplified model, we could argue for the substitution of the small resistors r by shortcircuits, arriving at the model in Figure 1.5. Another colleague, however, may argue equally forcefully that it is preferable to keep the small resistors r but, instead, to consider the large resistors R as open circuits, proposing the

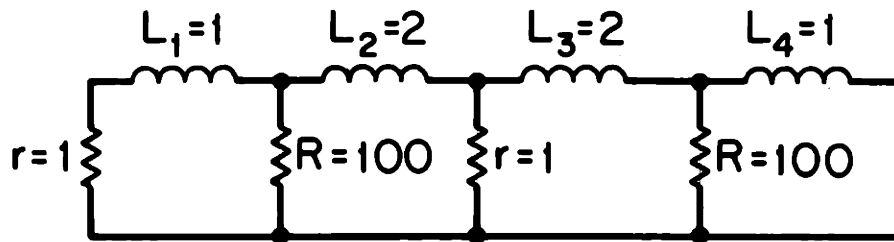


Figure 1.4: A network with fast and slow modes.

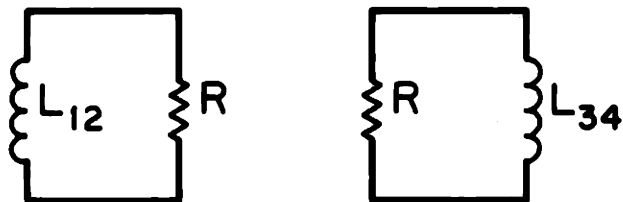


Figure 1.5: Approximation obtained by taking $r \approx 0$.

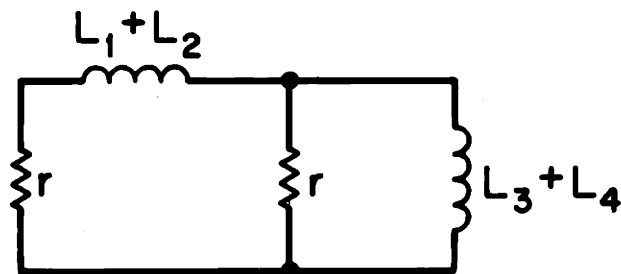


Figure 1.6: Approximation obtained by taking $R \approx \infty$.

circuit in Figure 1.6 as another approximate model. Who is right? The answer is, of course, nobody and everybody. The circuit in Fig. 1.5 gives a good approximation of the fast behavior of the network while the circuit in Fig. 1.6 is appropriate if the subject of interest is the behavior of the network on the long run. A complete picture of the different phenomena that occur in the network requires a combination of both reduced-order models.

The two examples just discussed can be fit into a single problem formulation which is the main object of study in this thesis. Specifically, we focus most of our attention on the asymptotic analysis of the system

$$\dot{x}^\varepsilon(t) = A(\varepsilon)x^\varepsilon(t) \quad (1.1.1)$$

(where $\varepsilon > 0$ is a small parameter) in the limit as $\varepsilon \downarrow 0$ and over the time interval $[0, \infty)$. The small parameter ε that enters in the system matrix models the presence of weak couplings between different variables of the system. The interval of study $[0, \infty)$ reflects our interest in all phenomena, fast or slow taking place in (1.1.1) and we seek aggregated models that are good approximations in the limit as $\varepsilon \downarrow 0$.

1.2 Description of Results

Motivated by simple examples such as those discussed in the preceding section, we have undertaken a detailed analysis of the equation:

$$\dot{x}^\varepsilon(t) = A_0(\varepsilon)x^\varepsilon(t) \quad (1.2.1)$$

or, equivalently, of $\exp\{A_0(\varepsilon)t\}$ for

$$A_0(\varepsilon) = \sum_{k=0}^{\infty} \varepsilon^k A_{k0} \quad (1.2.2)$$

and $\varepsilon \in [0, \varepsilon_0]$ a small parameter.

The central result in this thesis - a fundamental theoretical result which constitutes the foundation for the entire thesis - is a uniform asymptotic approximation of $\exp\{A(\varepsilon)t\}$ that we obtain under a certain *multiple semistability condition* (Theorem 4.4.4). In essence, the result states, in a precise way, that:

$$\exp\{A_0(\varepsilon)t\} \approx \exp\{A_m \varepsilon^m t\} \exp\{A_{m-1} \varepsilon^{m-1} t\} \dots \exp\{A_1 \varepsilon t\} \exp\{A_0 t\} \quad (1.2.3)$$

where the A_k are matrices whose computation we describe. Equation (1.2.3) explicitly decomposes $\exp\{A_0(\varepsilon)t\}$ into a set of evolutions taking place at different time scales. Each of the matrices A_k determines a reduced-order* model of (1.2.1) and the combined dimensionality of these models equals the dimension of the original system.

* In the control literature the expression "reduced-order models" is often preferred to "aggregated models". The former refers to any model simplification that involves a reduction in dimensionality while the latter (almost exclusively used in the economics literature) is usually restricted to a special type of reduced-order models in which the variables of the simplified model are obtained by simply adding groups of variables of the original system. We use them interchangeably.

The innovative aspects of this formulation and of the result mentioned are threefold:

- i) We do not require that the state variables in (1.2.1) be chosen so as to display the time scale structure of the system; rather the determination of this structure is one of the *outputs* of our construction.
- ii) We can handle systems with multiple time scales ($m > 1$), and we obtain a *uniform* asymptotic expansion on $[0, \infty)$.
- iii) We give an algorithm to compute the set of increasingly simplified reduced-order models valid at progressively slower time scales.

From a mathematical point of view, this is a result on singularly perturbed ordinary differential equations, its novel aspect being the uniform nature of the asymptotic approximation over the infinite time interval $[0, \infty)$, as opposed to pointwise approximations that have typically been used. From this point of view, we prove that (1.2.1) is singularly perturbed if and only if the rank of $A_0(\epsilon)$ changes at $\epsilon=0$, and we show that only singular perturbations result in multiple time scale behavior.

Further analysis of the significance of the ϵ -dependence in (1.2.1) for modelling purposes shows that in addition to time scale separation, models of this type can also exhibit unbounded amplitudes

as $\varepsilon \downarrow 0$. In this respect, we show that the multiple semistability condition referred to previously is indeed necessary and sufficient for the ε -dependence in (1.2.1) to model time scale separation exclusively (Theorem 4.4.9) and that only under this condition is it possible to perform a complete time scale analysis of (1.2.1). A brief consideration of the difficulties encountered when this condition is violated closes our analysis of the general LTI system case.

An application of particular interest to us has been the aggregation of finite state Markov processes (FSMP's) with rare transitions (modelled, as in the simple example given in Section 1.1, by small transition rates). If $A_0(\varepsilon)$ denotes now the matrix of transition rates of a FSMP $\eta^\varepsilon(t)$, its matrix of transition probabilities is given by

$$p^\varepsilon(t) = \exp\{A_0(\varepsilon)t\} \quad (1.2.4)$$

and thus all results mentioned for the general LTI system case have a direct interpretation for FSMP's. Indeed, stronger results are found to hold in this particular case: every singularly perturbed matrix of transition rates satisfies the multiple semistability condition and *every singularly perturbed FSMP can be hierarchically aggregated*. (Theorem 5.5.3). That is, the reduced-order models A_k that appear in the approximation (1.2.3) can be interpreted as the transition rates of a FSMP obtained by collapsing several states of

the original process $\eta^E(t)$ into one state of the aggregated model. The sequence of models A_k , $k=0,1,\dots,m$, forms a hierarchy because each of the aggregated models can be obtained by collapsing some states of the (already aggregated) model that precedes it. These models are adequate to classify events in $\eta^E(t)$ according to their rarity and they describe changes in $\eta^E(t)$ with an increasingly lower degree of detail.

The interpretation of limiting results and aggregation for FSMP's is shown to be greatly facilitated by the introduction and detailed consideration of *stochastically discontinuous* processes (i.e., processes with instantaneous transitions). As part of our work on FSMP's the usual continuity conditions imposed on FSMP's have been relaxed to allow for instantaneous transitions, and a full analysis of the properties of such processes has been conducted. As a result, we have established that in the general case the matrix of transition probabilities of a FSMP is of the form:

$$P(t) = \Pi \exp\{At\} \tag{1.2.5}$$

where Π is a projection (in contrast to $P(t) = \exp\{At\}$ for the usual stochastically continuous case) and that every stochastically discontinuous FSMP is uniquely determined by its ergodic projection at zero, Π , and an aggregated version of the process that is stochastically continuous. Specifically,

$$p(t) = \Pi \exp\{At\} = v \exp\{UAVt\}U \tag{1.2.6}$$

where U and V are certain aggregation matrices that are obtained by decomposing $\Pi=V \cdot U$ and UAV is the matrix of transition rates of the aggregated version (Theorem 5.3.5).

Another area in which an aggregation operation based on time scale separation seems to be possible and useful, the use of coherence area models for electric power systems, has been explored. We have shown that if a small parameter ϵ is introduced in the approximated version of the linearized swing equations used in power systems analysis as indicated below:

$$\dot{x}^{\epsilon}(t) = \begin{bmatrix} 0 & A(\epsilon) \\ I & 0 \end{bmatrix} x^{\epsilon}(t) \quad (1.2.7)$$

then an aggregated model of (1.2.7) can also be constructed and this model can be interpreted as replacing groups of generators of the original system (i.e., a coherent area) by single equivalent generator, which is precisely the heuristic way in which coherence area approximations are thought of and used in practice. Because under certain conditions, usually met in practice, the matrix $A(\epsilon)$ is also the matrix of transition rates of a FSMP, the coherence area aggregation of (1.2.7) is found to be completely equivalent to the hierarchical aggregation of FSMP's.

In addition to interpreting the aggregated model of (1.2.7) we have also indicated the nature of the approximation involved. The system (1.2.7) violates a basic condition required for the general result (1.2.3) to hold and thus the coherence area approximation is *not a uniform* asymptotic approximation. Rather, it must be interpreted on an entry by entry basis, and the interval of validity is different for different entries. Aside from the relevance of these results for the coherence area problem, they are also important because they give an indication of how (1.2.3) should be modified to accommodate systems in which the ϵ -dependence produces high-amplitudes in addition to time scale separation.

We have also explored the use of aggregated models in the construction of hierarchically structured suboptimal filters for singularly perturbed FSMP's through a simple example, we have introduced several qualitative performance measures for such filters and we have analyzed the trade-off between detection delays and the occurrence of false alarms in the detection of rare events. We have also shown how to build a detector for rare transitions that gives a correct reading "most of the time" (a new performance criterion that we introduce). For filtering problems with observations that only convey information about slow, aggregate changes we have indicated how to construct a filter using an aggregate model of the process to be estimated, and we have shown that it must be near-optimal "most of the time."

Finally, we discuss hierarchical filters with two channels of information (aggregate and decentralized measurements), mainly in heuristic terms, with the objective of suggesting further work along these lines. Several conjectures are stated for that purpose in the final pages of Chapter VII.

1.3 Thesis Outline

The thesis is organized in eight chapters. Following this introductory chapter we present a survey of relevant literature in Chapter II. Work that is related to our research is mentioned and briefly commented upon. Several specific references mentioned in Chapter II are reviewed with more detail in Chapters IV and V where they are contrasted with our results. For the reader's convenience we have put together in Chapter III a summary of the basic mathematical tools that we use in the rest of the thesis: Perturbation Theory for Finite Dimensional Linear Operators. Also included in this chapter the reader will find the definitions and notation we use for asymptotic analysis and a brief development of several non-standard topics in linear dynamical systems.

Chapter IV contains our basic results on the asymptotic analysis of LTI systems. The definition of time-scale behavior is introduced, the importance of the multiple semistability condition is established and the basic theorems on the asymptotic behavior of singularly

perturbed LTI systems are proven. In Chapter V FSMP's with rare transitions are studied in detail. It is shown that this class of models is especially suited for the multiple time-scale analysis developed in Chapter IV because the technical conditions required for the general results to hold are found to be satisfied always for FSMP's. As a tool in the development of aggregation techniques for FSMP's and also as an interesting result in itself, we present in Chapter V a complete analysis of stochastically discontinuous FSMP's.

Chapter VI deals with the multiple time scale analysis of the "swing" equations for electrical power networks and shows how the notion of coherence areas can be made precise our results on aggregation of FSMP's are crucial here. Chapter VII is of a more exploratory nature and represents an attempt to use aggregated models to simplify filtering problems for FSMP's. Several initial results and specific open questions are presented.

Finally, in Chapter VIII we summarize the main contributions of this thesis and suggest possible research directions motivated by our work.

CHAPTER II: PREVIOUS WORK AND RELATED LITERATURE

2.1 Introduction

A large number of researchers have contributed to the development and application of different aspects of the aggregation philosophy exposed in Chapter I and a comprehensive survey of all this work is beyond the scope of this thesis. In this chapter we will refer, however, to several survey papers and books for the reader interested in exploring the field in depth. The contributions and applications are scattered throughout several disciplines: econometrics [Chi 76], operation research [Zip 80], control theory [Kok 76], chemical engineering [Hil 77], stochastic processes [Cur 77], differential equations [Hop 71], linear algebra [Vis 60], functional analysis [Kor 78], etc., and, as a result, they are difficult to master and be put into context by a single researcher. Here we focus on the areas specifically relevant to the work described in this thesis.

Our purpose in this chapter is to provide an overview of a variety of results in the literature. In subsequent chapters we will discuss in more detail specific results related to our work in order to establish a clear picture of how our results relate to previous developments. In Section 2.2 we review previous work in the general area of singularly perturbed differential equations. In Section 2.3 we discuss several investigations on aggregation of Markov processes and finally in Section 2.4

we briefly mention several applications of these ideas to filtering and control problems.

2.2 Asymptotic Analysis of Singularly Perturbed Differential Equations

In the early stages of research into the properties of differential equations, it was recognized that exact closed-form solutions are possible only for a limited number of special equations and therefore the question of approximate solutions was immediately considered as a main area of research. Until the emergence of the digital computer (which made possible numerical methods), the most popular methods were asymptotic methods. The object of study was assumed to be a differential equation of the form:

$$\dot{x}^\varepsilon(t) = f(x^\varepsilon(t), t, \varepsilon) \quad (2.2.1)$$

where $x^\varepsilon(t)$ is a vector and ε a small parameter. What is meant by the term "asymptotic solution" is a function $\tilde{x}(t, \varepsilon)$ such that the difference $x(t, \varepsilon) - \tilde{x}(t, \varepsilon)$ is small (in some norm) in a given domain of interest. Furthermore, in this specific context what was sought were asymptotic solutions that could be expressed as combinations of solutions to equations that could be solved in terms of elementary functions.

Originally, equation (2.2.1) was exclusively studied over the interval $[0, T]$ and it was soon established that if the right hand side depends continuously on ε , then the solution also depends continuously in ε , which is a nice but not terribly interesting feature from a mathematical viewpoint. Motivated by problems in fluid dynamics,

non-linear oscillations, etc., Wasow and Tihonov (see [Was 44], [Was 65], [Vas 67] and references therein) shifted the attention of research to the case where the right hand side of (2.2.1) is singular at $\epsilon=0$. Specifically, they pointed out that a number of physically motivated problems reduce to differential equations containing small parameters multiplying the derivatives as follows:

$$\begin{aligned}\dot{x}(t) &= F(x(t), y(t), t, \epsilon) \\ \epsilon \dot{y}(t) &= G(x(t), y(t), t, \epsilon)\end{aligned}\tag{2.2.2}$$

and the study of this pair of differential equations became the paradigm that for nearly 50 years has dominated research in the area of singularly perturbed o.d.e.'s. Literally thousands of papers have been written on this subject and the reader is referred to [Vas 63,76, 78], [But 70], [Vis 62], [Hop 71], [Lag 72], for comprehensive surveys and to [Was 65], [Nay 73], [Ben 78] and [Eck 79] for more pedagogical, book-length expositions.

For our purposes, we want to emphasize two main limitations of the formulation (2.2.2). First, it *assumes* that a partition of the state space between the fast components (i.e., $y(t)$) and the slow components (the vector $x(t)$) is readily available, and, second, by studying the equation for $t \in [0, T]$, it singles out a specific time scale of interest neglecting the evolution of the system at slower time scale ($\tau = t/\epsilon$ for example) or faster time scales ($\tau' = t\epsilon$ for example). The result generally sought is a differential equation of the form:

$$\tilde{\mathbf{x}}(t) = \tilde{\mathbf{F}}(\tilde{\mathbf{x}}(t), t) \quad (2.2.3)$$

that approximates the behavior of the slow component $\mathbf{x}(t)$ and this can be considered a technique for order-reduction based on time-scale separation but with the limitations mentioned above.

Hoppensteadt recognized the need to go beyond the bounded interval $[0, T]$ and in [Hop 66] and [Hop 71] he studied the behavior of (2.2.2) over the interval $t \in [0, \infty)$. Under certain stability conditions he constructed an asymptotic approximation to the solution of (2.2.2) by combining the solution of the following two reduced-order problems:

Slow time scale

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), \mathbf{y}(\mathbf{x}(t)), t, 0) \\ 0 = \mathbf{G}(\mathbf{x}(t), \mathbf{y}(\mathbf{x}(t)), t, 0) \end{cases} \quad (2.2.4)$$

Fast time scale

$$\begin{cases} \dot{\mathbf{x}}(t) = 0 \\ \dot{\mathbf{y}}(t) = \mathbf{G}(\mathbf{x}, \mathbf{y}(t), t, 0) \end{cases} \quad (2.2.5)$$

The uniform asymptotic approximation involves a combination of a fast component $\tilde{\mathbf{y}}(t)$ and a slow component $\tilde{\mathbf{x}}(\varepsilon t)$ that only evolves at the slow time scale t/ε . This work clearly indicated the need to use more and one reduced-order model of a system (equations (2.2.4) and (2.2.5) in this case) to produce a good approximation of all the features of the system evolution.

The inherent difficulty of nonlinear equations and the richness of results obtained using the formulation (2.2.2)* kept researchers away from examining problems involving more than two time scales and also away from formulations which do not assume an a priori known separation between slow and fast states. With hindsight it seems rather strange that, as a first step towards the understanding of (2.2.1), the linear, constant coefficient case

$$\dot{x}(t) = A(\epsilon)x(t) \tag{2.2.6}$$

was not studied in detail. One must remember, however, the original motivation for asymptotic methods. Equation (2.2.6) can in principle be solved *exactly* for any value of ϵ , in terms of well known exponential functions and therefore, from the point of view of the early investigations, no asymptotic approximations were required. It took again another class of problems with no closed-form solution, evolution equations in abstract Banach (i.e., infinite dimensional) spaces to focus attention on equation (2.2.6), and it was also Hoppensteadt who took the lead in [Hop 75] studying the asymptotic behavior of

$$\dot{x}^\epsilon(t) = A(\epsilon)x^\epsilon(t) + G(x^\epsilon(t), \epsilon) \tag{2.2.7}$$

* The results of Hoppensteadt refer only to the case where the equation $G(x, y, t) = 0$ has a unique solution $y(x)$. A wealth of interesting cases occur if several isolated solutions are possible (bifurcation theory) or a continuum of solutions exist ([Vas 78]). Accordingly, most of the research was directed towards those problems rather than reconsidering the formulation (2.2.2).

over the time interval $t \in [0, \infty)$. (For purposes of comparison to our work can set $G(x, \epsilon) \equiv 0$).

Hoppensteadt recognized that even if $A(\epsilon)$ depends continuously on ϵ , the solution $x^\epsilon(t)$ may actually depend on ϵ in a quite singular way if analyzed over the infinite time interval and he studied several examples where a uniform asymptotic approximation can be constructed as a (finite) sum of functions, each depending regularly on a different time scale, and each approaching a steady state as its time scale approaches infinity. For his development, he *assumed* that a decomposition of $A(\epsilon)$ in the form:

$$A(\epsilon) = A_0(\epsilon) \oplus \epsilon A_1(\epsilon) \oplus \dots \oplus \epsilon^m A_m(\epsilon) \quad (2.2.8)$$

was available with $A_i(\epsilon)$ continuous in ϵ and stable, and he showed that a uniform asymptotic approximation of $x^\epsilon(t)$ can be obtained as a linear combination of the solutions to the reduced-order models:

$$\begin{aligned} \dot{x}_0(t) &= A_0(0)x_0(t) \\ \dot{x}_1(t) &= \epsilon A_1(0)x_1(t) \\ &\vdots \\ \dot{x}_m(t) &= \epsilon^m A_m(0)x_m(t) \end{aligned} \quad (2.2.9)$$

Hoppensteadt's work is significant in several respects:

- i) it points out the importance of considering approximation that are uniformly valid over the infinite time interval;

- ii) it points out that even if the right hand side of (2.2.9) depends continuously on ϵ , the perturbation may have dramatic effects on $x^\epsilon(t)$ if the whole interval $[0, \infty)$ is considered; and
- iii) it shows that in general several time scales are needed to describe completely the asymptotic behavior of a perturbed linear system.

The direct application of Hoppensteadt's work to the decomposition and aggregation problem described in Chapter I is not possible, however, because his starting point (equation (2.2.8)) requires that the designer knows a priori how the system should be decomposed. From this perspective, our work can be seen as providing a systematic procedure for obtaining a decomposition as in (2.2.8) starting from an arbitrary $A(\epsilon)$.

The asymptotic analysis of abstract evolution equations without the a priori decomposition assumption has been carried out by several authors (see [Pap 75], [Dav 77,80] and [Kor 78]). In all cases, however, at most two time scales have been considered and no uniform approximation valid over $[0, \infty)$ has been sought. For the finite dimensional case that is the subject of this thesis, the work of Campbell and Rose (see [Cam 78a,78b,79,80]) represents a first attempt to analyze the equation

$$\dot{x}^\epsilon(t) = A(\epsilon)x^\epsilon(t) \quad (2.2.10)$$

without a priori assumptions on the structure of $A(\epsilon)$. These authors give several conditions under which the limits

$$\lim_{\epsilon \downarrow 0} x^\epsilon(t/\epsilon^r) \quad (2.2.11)$$

exist *pointwise* (i.e., for a given t , not uniformly in t) for different r , and they also give an expression for the limits. We discuss their work in more detail in Section 4.3 where we will contrast it with ours. In general terms, we can say that although the limits (2.2.11) can be interpreted as a time scale characterization of the system (2.2.10), the treatments in [Cam 78a,78b,79,80] do not address the question of how many and which time scales a system has and, more fundamentally, the question of asymptotic approximation or, equivalently, the question of how to combine the behavior of the system at several time scales to produce an approximation that is valid over the infinite time interval $[0, \infty)$.

In summary, from the viewpoint taken in this thesis there are two respects in which the existing literature on singularly perturbed o.d.e.'s leaves important questions partially answered: (1) in the problem formulation, in particular with regard to the assumption of a priori known time scale decomposition; and (2) in the interpretation of the results in terms of uniform asymptotic approximations combining behavior at several time scales. Our work is an attempt to answer these questions for finite-dimensional linear systems and to provide some insight into the problem of asymptotic approximations, by revealing the underlying pattern when multiple time scale behavior is present and by clarifying the intimate relationship between weak couplings, time scales, reduced-order modelling and asymptotic approximations.

2.3 Aggregation of Markov Processes

When in 1961 Simon and Ando ([Sim 61]) first made explicit the reasoning behind approximations based on aggregation ideas, they chose a Markov chain example to clarify their presentation. The association of each variable with a state of the process and the use of diagrams to visualize transitions between states make the idea of aggregation very intuitively appealing as a way to obtain coarse descriptions for such models. Furthermore, for reasons not clear then (which will be made clear to the reader in Chapter V) it was noticed that systems described by stochastic matrices were easier to aggregate than other linear systems. It is thus not surprising that most of the theoretical results on aggregation have been obtained by researchers working in one way or another with stochastic processes.

The first rigorous result was presented in [Kor 69] where it was shown that a Semi-Markov process with a matrix of transition probabilities of the form $P(\epsilon) = P_0 + \epsilon Q$ with $P_0 = \text{diag}\{P_{01}, P_{02}, \dots, P_{0m}\}$ (i.e., diagonally dominant) could be aggregated by collapsing states belonging to a group determined by, say, P_{0j} into a single state of the aggregated process. If the system under consideration is aggregated according to this partition then in the limit as $\epsilon \downarrow 0$ the transitions between *groups* of states follow a markovian law with aggregated parameters that can be computed from the detailed model of

the process. Different versions of the same basic idea were repeatedly presented in several papers by Korolyuk and coworkers (see [Kor 70], [Gus 71], [Tur 71] and [Kor 76,78]), and also in [Per 74], [Gai 75] and [Cur 77].

As in the literature on singular perturbations, no attempt was made in the literature cited above to analyze the general case where the matrix of transition probabilities (or transition rates for continuous time processes), $P(\epsilon)$, is not assumed to have the special diagonally dominant structure. As a result, only one aggregated model of a system was considered and only one time scale (besides the natural one) was ever considered. The idea of a hierarchy of models tied to an asymptotic approximation of the process is not present in the works mentioned above.

Only recently in [Cas 80] and [Del 82] the existence of a hierarchy of models in the general, not diagonally dominant case has been suggested. These works, whose relationship to our results is discussed in more detail in Section 5.2 stop short of providing a clear interpretation of the aggregated models nor do they make a rigorous connection between the aggregated models and an asymptotic approximation of the original process which is valid on the infinite time interval $[0, \infty)$. These are precisely the motivations behind our work presented in Chapter V.

2.4 Order-Reduction and Aggregation Methods in Filtering and Control Theory

The aggregation technique discussed in Chapter I is mainly a methodology for model simplification. Such methodologies are of practical significance to the extent that they are useful in simplifying the solution of problems posed for the detailed model of the system. Ideally, an aggregation methodology such as the one attempted in this thesis, in addition to showing how to decompose a model for a given problem, should also indicate how to define a set of problems posed for the aggregated models whose solutions can then be used to approximate the solution to the original problem.

In this thesis only minor attention has been given to this aspect of aggregation which is treated in Chapter VII in connection with some filtering problems. Here we briefly mentioned the use of aggregated models by other researchers in control and filtering applications.

The introduction of singular perturbation techniques in the area of filtering and control is due to Kokotovic and Haddad (see [Kok 70] and [Had 71]). A two time scale system represented by the pair of equations

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{F}(\mathbf{x}(t), \mathbf{y}(t), t, \varepsilon) \\ \varepsilon \dot{\mathbf{y}}(t) &= \mathbf{G}(\mathbf{x}(t), \mathbf{y}(t), t, \varepsilon)\end{aligned}$$

was introduced to justify the classical "dominant mode" techniques which neglect "high-frequency" parts and retain "low-frequency" components of models. The pioneering work of Kokotovic and Haddad signaled a new approach to modelling that has since then been the subject of intense research (see for example the classic survey [Kok 76] and the collection of papers in [Kok 80]). The standard model for linear systems with two time scales that was popularized by these authors has the following form:

$$\begin{cases} \dot{x}_1(t) = A_{11}x_1(t) + A_{12}x_2(t) \\ \epsilon \dot{x}_2(t) = A_{21}x_1(t) + A_{22}x_2(t) \end{cases} \quad (2.3.1)$$

and from it to reduced-order models are obtained:

Slow model

$$\begin{aligned} \dot{x}_1(t) &= (A_{11} - A_{12}A_{22}^{-1}A_{21})x_1(t) \\ x_2(t) &= -A_{22}^{-1}A_{21}x_1(t) \end{aligned} \quad (2.3.2)$$

Fast model

$$\begin{aligned} \dot{x}_1(t) &= 0 \\ \dot{x}_2(t) &= A_{22}x_2(t) \end{aligned} \quad (2.3.3)$$

Under the conditions usually assumed (A_{22} and $A_{22}A_{11}^{-1}A_{12}^{-1}A_{21}$ stable), the system (2.3.1) exhibits only two time scales and (2.3.2) and (2.3.3) are reduced-order models that adequately describe the slow and fast behavior respectively.

Filtering and control problems posed for (2.3.1) have been successfully decomposed into separate problems posed for (2.3.2) and (2.3.3) (see for example [Kok 72], [Phi 80], [Had 77], [Ten 77]). In practical situations, however, the a priori decomposition assumed in (2.3.1) has been a source of difficulties when trying to apply these results. Attempts to decomposed models of electric power systems in the canonical form (2.3.1) have shown that this is by no mean a simple matter (see [Avr 80] and [Win 80]). Ad-hoc algorithms have been proposed but definite results are not available. We think that because our formulation does not require the model to have any specific structure, it may facilitate the application of theoretical results to practical systems. We, however, do not deal with this problem here.

The control of Markov chains with a decomposable structure has also been studied by several researchers recently (see [Ten 80], [Del 81] and [Phi 81]). A diagonally-dominant structure has routinely been assumed in these papers and the time scale separation has been shown to lead to hierarchical algorithms in which fast subsystem optimizations are coordinated at a slower aggregated level. Our results on FSMP's aggregation should facilitate the decomposition of control

algorithms for general, not necessarily diagonally-dominant processes and should give rise to hierarchical optimization algorithms with a multilevel hierarchy. We comment more on this in Chapter VII when discussing filtering problems for FSMP's.

CHAPTER III: MATHEMATICAL PRELIMINARIES

3.1 Introduction

This chapter establishes notation and nomenclature while providing the reader with a concise presentation of several results that constitute the starting point for the theory developed in latter chapters. Most of these results are available in the literature; proofs are given only for those which are not readily available in the desired form. Basic results on linear operators and perturbation theory in finite dimensional spaces are reviewed in Sections 3.2 and 3.4 following [Kat 66]. The definitions and notation for asymptotic analysis of singular perturbations introduced in Section 3.3 are those of [Eck 79]. Section 3.5 contains standard definitions and results in dynamical systems and positive linear systems as given for example in [Lue 79].

3.2 Linear Operators in Finite Dimensional Spaces

We assume that the reader is familiar with elementary notions of linear algebra and analysis although some fundamental results are collected for the convenience of later reference. The eigenvalue problem is dealt with in more detail since it is extensively used in subsequent chapters. The approach is analytic rather than algebraic, depending on a function-theoretic treatment of the resolvent.

3.2.1 Normed Vector Spaces

Let V be an n -dimensional vector space. For a fixed basis $\{v_j\}$ and $u \in V$, denote by $\|u\|$ the ℓ^∞ norm of u , i.e., if $u = \sum_{j=1}^n \alpha_j v_j$, $\alpha_j \in \mathbb{C}$ then

$$\|u\| = \max_j |\alpha_j| \tag{3.2.1}$$

With this definition V becomes a normed vector space. A sequence

$\{u_n\}_{n \geq 0}$ of vectors in V is said to converge to u , and we write $u_n \rightarrow u$, or

$$\lim_{n \rightarrow \infty} u_n = u, \text{ if } \lim_{n \rightarrow \infty} \|u_n - u\| = 0 \tag{3.2.2}$$

The convergence of an infinite series $\sum_{n=1}^{\infty} u_n$; the continuity

of a function $u(t)$ defined for a real or complex variable t and taking values in V ; and the derivative and the Riemann or contour integrals of $u(t)$, are all defined as in the scalar case.

When $u(t)$ is defined and differentiable everywhere in a domain D of the complex plane, $u(t)$ is said to be *analytic (regular) or holomorphic* in D . Throughout this thesis we use standard definitions and results of complex function theory such as Cauchy's integral theorem, Laurent's expansions, etc. as given, for example, in [Kno 45]. We also apply these results to vector- or operator-valued functions without particular comments.

3.2.2 Linear Operators

Let V and W be two vector spaces and $T:V \rightarrow W$ a *linear operator*. The image of V under T is called the range of T and is denoted by $R(T)$. The dimension of $R(T)$ is called the rank of T ; we denote it by $\text{rank } T$. The inverse image of the zero element of W is called the *null space* of T and is denoted by $N(T)$. The dimension of $N(T)$ is called the *nullity* of T which we denote by $\text{nul } T$. A basic result in linear algebra is:

$$\text{rank } T + \text{nul } T = \dim V \tag{3.2.3}$$

If T maps V onto W in a one to one fashion, the inverse operator $T^{-1}: W \rightarrow V$ is well defined and T is said to be *non-singular*, otherwise is said to be *singular*. With the obvious definitions for addition, scalar multiplication and product, the set of linear operator on V to itself is an *algebra* (non-commutative if $\dim V \geq 2$). I and 0 will denote the identity and the zero elements respectively.

If $\dim V = n$ and $\dim W = m$ then T can be represented by an $(m \times n)$ matrix for given bases for V and W . Conversely, any $(m \times n)$ matrix determines a linear operator with respect to the given basis. We will often work with the matrix representation of linear operators.

Let X and Y be two subspaces of V such that each $u \in V$ can be uniquely decomposed in the form $u = u' + u''$ with $u' \in X$ and $u'' \in Y$, i.e., $V = X \oplus Y$. The linear operator $P: V \rightarrow V$, $Pu = u'$ is called the *projection on X along Y* and we have $R(P) = X$, $N(P) = Y$. P is *idempotent*, i.e., $P^2 = P$ and conversely any idempotent operator is a projection. More generally

$$V = X_1 \oplus \dots \oplus X_s \tag{3.2.4}$$

$u = u_1 + \dots + u_s$, $u_i \in X_i$ and the operator P_j defined by $P_j u = u_j$ is the projection on X_j along $X_1 \oplus \dots \oplus X_{j-1} \oplus X_{j+1} \oplus \dots \oplus X_s$. Furthermore, we have

$$\sum_{j=1}^s P_j = I \tag{3.2.5}$$

$$P_j P_k = \delta_{kj} P_j \tag{3.2.6}$$

Conversely, any set of operators P_j satisfying (3.2.5) and (3.2.6) is a family of projections that determine the direct sum decomposition (3.2.4) with $X_i = R(P_j)$. A basis $\{v_j\}$ of V is said to be *adapted* to the decomposition (3.2.4) if the first $n_1 = \dim X_1$ elements of $\{v_j\}$ belong to X_1 , the following $n_2 = \dim X_2$ ones belong to X_2 and so on.

A subspace X of V is said to be *invariant* under a linear operator $T: V \rightarrow V$ if $TX \subset X$. In this case T induces a linear operator $T: X \rightarrow X$ defined by $T_x u = Tu$ for $u \in X$ which is called the *part* of T in X . T is said to be *decomposed* by a set of subspaces $\{X_i\}$ if (3.2.4) is satisfied and all the X_i are invariant under T . In this case T is completely determined by its parts T_{X_i} and if $\{P_j\}$ is the set of projections corresponding to (3.2.4), then T commutes with each P_i . Conversely, T is decomposed by $\{X_j\}$ if it commutes with all the P_j . The operator $P_i T = T P_i = P_i T P_i$ coincides with T and with T_{X_i} when applied to $u \in X_j$; it is sometimes identified with T_{X_i} when no confusion is possible.

A linear operator $T: V \rightarrow V$ is called *nilpotent* if $T^r = 0$ for some positive integer r . A nilpotent operator is necessarily singular.

For future reference we will need the following lemma on the full-rank factorization of singular matrices.

Lemma 3.2.1

- i) Let T be an $(n \times n)$ matrix of rank $r \leq n$. Then there exist matrices F and G , $(n \times r)$ and $(r \times n)$ respectively, of rank r , such that

$$T = F.G$$

ii) Let the square matrix T have full rank factorization

$$T = F \cdot G$$

Then T is a projection if and only if

$$F \cdot G = I$$

Proof: See [Ben 80], pages 22 and 49. □

3.2.3 Analysis with Operators

The set of all linear operators on V to W is a normed vector space with norm induced by the vector norms in V and W as follows:

$$\|T\| = \sup_{\substack{u \in V \\ u \neq 0}} \frac{\|Tu\|}{\|u\|} = \sup_{\substack{u \in V \\ \|u\|=1}} \|Tu\| \quad (3.2.7)$$

If (τ_{ij}) is the matrix representation of T in the given basis, then

$$\|T\| = \max_i \sum_j |\tau_{ij}| \quad (3.2.8)$$

As a feature of the induced norm, note that

$$\|TS\| \leq \|T\| \cdot \|S\| \quad (3.2.9)$$

The convergence of a sequence of operators, $\{T_n\}$ and that of an infinite series of operators $\sum_{n=0}^{\infty} T_n$ is defined as in the vector or scalar

cases. Similarly, the *absolute convergence* of such series means that the series $\sum_{n=0}^{\infty} \|T_n\|$ is convergent. In this case $\sum_{n=0}^{\infty} T_n$ is convergent with

$$\left\| \sum_{n=0}^{\infty} T_n \right\| \leq \sum_{n=0}^{\infty} \|T_n\|.$$

Example 3.2.2: Exponential function

$$\exp\{Tt\} \triangleq \sum_{n=0}^{\infty} \frac{1}{n!} T^n t^n \quad (3.2.10)$$

This series is absolutely convergent for any complex number t because

$$\left\| \frac{1}{n!} T^n t^n \right\| \leq \frac{1}{n!} \|T\|^n |t|^n \quad \text{and we have}$$

$$\|\exp\{Tt\}\| \leq \exp\{\|T\| \cdot |t|\} \quad (3.2.11)$$

□

Example 3.2.3: Neumann series

$$(I-T)^{-1} = \sum_{n=0}^{\infty} T^n \quad (3.2.12)$$

This series is absolutely convergent if $\|T\| < 1$. In this case

$$\|(I-T)^{-1}\| \leq (1 - \|T\|)^{-1} \quad (3.2.13)$$

□

Operator-valued functions $T(t)$ defined for a real or complex variable t (such as (3.2.10) for example) can be defined and treated as vector-valued or scalar functions. The following lemma dealing with projection-valued functions will be useful in latter chapters.

Lemma 3.2.4: ([Kat 66] p.34)

Let $P(t)$ be a projection depending continuously on a parameter t varying in a (connected) region of real or complex numbers. Then the ranges $R(P(t))$ for different t are isomorphic to one another. In particular, $\dim R(P(t))$ is constant.

□

3.2.4 The Resolvent

Let T be a linear operator on V to itself. A complex number λ is called an eigenvalue of T if there exists a non-zero vector u such that

$$Tu = \lambda u \tag{3.2.14}$$

u is called an *eigenvector* of T with eigenvalue λ . The subspace of eigenvectors of T with eigenvalue λ is called the *geometric eigenspace* for λ and its dimension the *geometric multiplicity* of λ . Eigenvectors with different eigenvalues are linearly independent and therefore there are at most $n = \dim V$ eigenvalues of T . The set of all eigenvalues of T is called the *spectrum* of T ; we denote it by $\sigma(T)$.

If X is an invariant subspace of T then any eigenvalue [eigenvector] of T_X (the part of T in X) is an eigenvalue [eigenvector] of T . An eigenvalue of T_X is called an *eigenvalue of T in X* . If P is a projection that commutes with T , T is decomposed according to $V = X \oplus Y$, $X = R(P)$, $Y = N(P)$. It is sometimes convenient to consider the eigenvalue problem for the parts of T in X and Y separately. As said before, the part of T in X , T_X , may be identified with $TP = PT = PTP$. It should be noticed, however, that if T_X has an eigenvalue zero with geometric eigenspace L_0 then the geometric eigenspace of the zero eigenvalue of TP is $Y + L_0$.

The operator-valued function

$$R(\xi, T) = (T - \xi I)^{-1} \tag{3.2.15}$$

is well defined for any complex number $\xi \in \rho(T) = C - \sigma(T)$ and it is called the *resolvent* of T . The set $\rho(T)$ is referred to as the *resolvent set* of T .

$R(\xi, T)$ satisfies the so called *resolvent equation*

$$R(\xi_1, T) - R(\xi_2, T) = (\xi_1 - \xi_2) R(\xi_1, T) R(\xi_2, T) \quad (3.2.16)$$

which, in particular, implies that $R(\xi_1, T)$ and $R(\xi_2, T)$ commute. The resolvent is an analytic function with isolated singularities at precisely the eigenvalues λ_k , $k=0, 1, \dots, s$, of T .

The Laurent series of $R(\xi, T)$ at λ_k has the form:

$$R(\xi, T) = -(\xi - \lambda_k)^{-1} P_k - \sum_{n=1}^{m_k - 1} (\xi - \lambda_k)^{-n-1} D_k^n + \sum_{n=0}^{\infty} (\xi - \lambda_k)^n S_k^{n+1} \quad (3.2.17)$$

where

$$P_k = -\frac{1}{2\pi i} \int_{\Gamma_k} R(\xi, T) d\xi \quad (3.2.18)$$

(with Γ_k a positively oriented contour enclosing λ_k but no other eigenvalue of T) is a projection called the *eigenprojection* for the eigenvalue λ_k of T ; $m_k = \dim R(P_k)$ is the *algebraic multiplicity* of λ_k ;

$$D_k = -\frac{1}{2\pi i} \int_{\Gamma_k} (\xi - \lambda_k) R(\xi, T) d\xi \quad (3.2.19)$$

is the *eigennilpotent* ($D_k m_k = 0$) for the eigenvalue λ_k of T ; and

$$S_k = \frac{1}{2\pi i} \int_{\Gamma_k} (\xi - \lambda_k)^{-1} R(\xi, T) d\xi \quad (3.2.20)$$

It is not difficult to see that the following relations hold:

$$P_k S_k = S_k P_k = 0 \quad (3.2.21)$$

$$P_k D_k = D_k P_k = D_k \quad (3.2.22)$$

$$P_k T = T P_k \quad (3.2.23)$$

$$(T - \lambda_k I) S_k = I - P_k \quad (3.2.24)$$

$$(T - \lambda_k I) P_k = D_k \quad (3.2.25)$$

$$P_k P_\ell = \delta_{k\ell} P_k \quad (3.2.26)$$

$$\sum_{k=1}^s P_k = I \quad (3.2.27)$$

From (3.2.26) and (3.2.27) it follows that

$$V = M_1 \oplus \dots \oplus M_s \quad (3.2.28)$$

with $M_k = R(P_k)$. Since the P_k commute with T and with one another, the M_k are invariant subspaces for T . M_k is called the *algebraic eigenspace* for the eigenvalue λ_k of T . It follows from (3.2.22) and (3.2.25) that

$$T P_k = P_k T = P_k T P_k = \lambda_k P_k + D_k \quad (3.2.29)$$

which together with (3.2.27) gives the *canonical form* or *spectral representation* of T :

$$T = \sum_{k=1}^s (\lambda_k P_k + D_k) \quad (3.2.30)$$

An eigenvalue λ_k is said to be *semisimple* if the associated eigen-nilpotent D_k is zero and *simple* if in addition $m_k=1$. T is said to be *diagonalizable* if all its eigenvalues are semisimple. If all eigenvalues of T are simple T is itself called *simple*; in this case T has n eigenvalues.

For any polynomial $p(z) = \alpha_0 + \alpha_1 z + \dots + \alpha_m z^m$ we define the operator

$$p(T) = \alpha_0 I + \alpha_1 T + \dots + \alpha_m T^m \quad (3.2.31)$$

Making use of the resolvent we can define more general functions of T as follows: Suppose that $\phi(\xi)$ is analytic in a domain Δ of the complex plane containing all the eigenvalues λ_k of T , and let $\Gamma \subset \Delta$ be a positively-oriented contour enclosing all λ_k in its interior. Then $\phi(T)$ is defined by:

$$\phi(T) = - \frac{1}{2\pi i} \int_{\Gamma} \phi(\xi) R(\xi, T) d\xi \quad (3.2.32)$$

This definition coincides with (3.2.31) when $\phi(T)$ is a polynomial and it provides a useful representation for the exponential function whose equivalent series definition was given in (3.2.10). That is,

$$\exp\{Tt\} = - \frac{1}{2\pi i} \int_{\Gamma} e^{\xi t} R(\xi, t) d\xi \quad (3.2.33)$$

If $T = \sum_{k=1}^s (\lambda_k P_k + D_k)$ then the spectral representation of $\phi(T)$ is

given by:

$$\phi(T) = \sum_{k=1}^s (\phi(\lambda_k) P_k + \tilde{D}_k) \quad (3.2.24)$$

where

$$\tilde{D}_k = \phi'(\lambda_k) D_k + \dots + \frac{\phi^{(m_k-1)}(\lambda_k)}{(m_k-1)!} D_k^{m_k-1} \quad (3.2.35)$$

are nilpotents commuting with each other and with the P_k .

3.2.5 Operators with semisimple null structure

In this section we analyze a class of linear operators that play a special role in the theory of aggregation developed in later chapters. The importance of this class of operators in aggregation problems is emphasized in [Kor 78], where the more general case of operators in infinite dimensional spaces is treated. We give here simplified proofs of the basic results for the finite-dimensional case.

Definition 3.2.5 A linear operator T_\circ on V to itself is said to have *semisimple null structure* (SSNS) if zero is a semisimple eigenvalue of T_\circ .

By convention we will refer to non-singular operators as having SSNS. The following lemma establishes some properties of operators with SSNS.

Lemma 3.2.5 The following are equivalent statements:

- i) T_\circ has SSNS
- ii) $V = R(T_\circ) \oplus N(T_\circ)$

$$\text{iii) } \mathcal{R}(T_{\circ}) = \mathcal{R}(T_{\circ}^2)$$

$$\text{iv) } \text{rank } T_{\circ} = \text{rank } T_{\circ}^2$$

$$\text{v) } \mathcal{N}(T_{\circ}) = \mathcal{N}(T_{\circ}^2)$$

Proof: i) \Rightarrow ii). Using the canonical form (3.2.30) for T_{\circ} with

$\lambda_{\circ} = 0$ we have $\sum_{k=1}^s (\lambda_k P_k + D_k)$ and $V = \sum_{k=0}^s \bigoplus \mathcal{R}(P_k)$. Clearly $\mathcal{R}(P_{\circ}) \subset \mathcal{N}(T_{\circ})$.

on the other hand, $T_{\circ}u = 0$ implies $\lambda_k u_k + D_k u_k = 0$ for $u_k = P_k u$, $k=1, \dots, s$ and

because D_k is nilpotent we have $u_k = 0$, $k=1, \dots, s$ giving $\mathcal{R}(P_{\circ}) = \mathcal{N}(T_{\circ})$.

Also, $\mathcal{R}(T_{\circ}) \subset \sum_{k=1}^s \bigoplus \mathcal{R}(P_k)$ and ii) follows from $\dim \mathcal{R}(T_{\circ}) = \dim \sum_{k=1}^s \bigoplus \mathcal{R}(P_k)$.

ii) \Rightarrow iii). Obviously $\mathcal{R}(T_{\circ}^2) \subset \mathcal{R}(T_{\circ})$ and $u = T_{\circ}z$ implies $u = T_{\circ}z'$, $z' \in \mathcal{R}(T_{\circ})$

thus $\mathcal{R}(T_{\circ}^2) = \mathcal{R}(T_{\circ})$.

iii) \Rightarrow iv) trivially.

iv) \Rightarrow v). Clearly $\mathcal{N}(T_{\circ}) \subset \mathcal{N}(T_{\circ}^2)$ and $\dim \mathcal{N}(T_{\circ}) = n - \text{rank } T_{\circ} = n - \text{rank } T_{\circ}^2 = \dim \mathcal{N}(T_{\circ}^2)$ thus giving v).

v) \Rightarrow i). Suppose $D_{\circ} \neq 0$, then $D_{\circ}^k \neq 0$ for $k=1, \dots, m-1$ and $D_{\circ}^m = 0$ for some integer $m > 1$. Choose v such that $D_{\circ}^{m-2}v \neq 0$ and $D_{\circ}^{m-1}v \neq 0$ and define $u = D_{\circ}^{m-2}v$. We then have $T_{\circ}u = D_{\circ}^{m-1}v \neq 0$ and $T_{\circ}^2u = D_{\circ}^m v = 0$, a contradiction. \square

It follows from the proof of ii) in Lemma 3.2.6 that if T_{\circ} has SSNS then P_{\circ} , the eigenprojection for the zero eigenvalue of T_{\circ} , is also the projection on $\mathcal{N}(T_{\circ})$ along $\mathcal{R}(T_{\circ})$. Let $Q_{\circ} = I - P_{\circ}$, then $P_{\circ}T_{\circ} = T_{\circ}P_{\circ} = 0$ and $Q_{\circ}T_{\circ} = T_{\circ}Q_{\circ} = T_{\circ}$.

Theorem 3.2.7 If T_0 has SSNS then $T_0 + P_0$ is non-singular.

Proof: Denote by \hat{T}_0 the restriction of T_0 on $R(T_0)$. \hat{T}_0 is invertible because $\hat{T}_0 R(T_0) = R(T_0)$. Let $u = u' + u''$ with $u' \in R(T_0)$, $u'' \in N(T_0)$ and define $\tilde{T}_u = \hat{T}_0^{-1} u' + u''$. Then $(T_0 + P_0) \tilde{T}_u = T_0 \hat{T}_0^{-1} u' + u'' = u$ and $\tilde{T}(T_0 + P_0)u = \hat{T}_0^{-1} T_0 u' + u'' = u$. Thus $\tilde{T} = (T_0 + P_0)^{-1}$. \square

Define the operator $T_0^\#$ by $T_0^\# = (T_0 + P_0)^{-1} - P_0$. The following lemma gives several properties of this operator.

Lemma 3.2.8

$$i) \quad P_0 T_0^\# = T_0^\# P_0 = 0 \quad (3.2.37)$$

$$ii) \quad Q_0 T_0^\# = T_0^\# Q_0 = T_0^\# \quad (3.2.38)$$

$$iii) \quad T_0^\# T_0 = T_0 T_0^\# = Q_0 \quad (3.2.39)$$

$$iv) \quad |||T_0^\#||| = |||\hat{T}_0^{-1}||| \quad (3.2.40)$$

Proof: Let $u = u' + u''$ with $u \in R(T_0)$ and $u' \in R(T_0)$. By definition $T_0^\# u = \hat{T}_0^{-1} u'$ and i) to iv) follow immediately.

It follows from lemma 2.3.8 that $T_0 T_0^\# T_0 = T_0$, $T_0^\# T_0 T_0^\# = T_0^\#$ and $T_0 T_0^\# = T_0 T_0^\# = T_0^\# T_0$. $T_0^\#$ is thus the group generalized inverse of T_0 (see [Cam 79a]); we will refer to it simply as the *generalized inverse* of T_0 . The following lemma shows that if T_0 has SSNS then P_0 and $T_0^\#$ fully determine the Laurent expansion of the resolvent $R(\lambda, T_0)$ at zero.

Lemma 3.2.9: If T_{\circ} has SSNS, then for $0 < |\lambda| < \|T_{\circ}^{\#}\|^{-1}$

we have

$$R(\lambda, T_{\circ}) = -\frac{P_{\circ}}{\lambda} + \sum_{k=0}^{\infty} \lambda^k (T_{\circ}^{\#})^{k+1} \quad (3.2.41)$$

$$= -\frac{P_{\circ}}{\lambda} + T_{\circ}^{\#} (I - \lambda T_{\circ}^{\#})^{-1} \quad (3.2.42)$$

Proof: For $0 < |\lambda| < \|T_{\circ}^{\#}\|^{-1}$ the above series is absolutely convergent and the equality between (3.2.41) and (3.2.42) follows from (3.2.12).

On the other hand, it follows from Lemma 3.2.8 that:

$$\begin{aligned} (T_{\circ} - \lambda I) \left(-\frac{P_{\circ}}{\lambda} + \sum_{k=0}^{\infty} \lambda^k (T_{\circ}^{\#})^{k+1} \right) &= \\ Q_{\circ} \sum_{k=0}^{\infty} \lambda^k (T_{\circ}^{\#})^k + P_{\circ} - \sum_{k=1}^{\infty} \lambda^k (T_{\circ}^{\#})^k &= \\ Q_{\circ} + \sum_{k=1}^{\infty} \lambda^k (T_{\circ}^{\#})^k + P_{\circ} - \sum_{k=1}^{\infty} \lambda^k (T_{\circ}^{\#})^k &= I \end{aligned}$$

and similarly $\left(-\frac{P_{\circ}}{\lambda} + \sum_{k=0}^{\infty} \lambda^k (T_{\circ}^{\#})^{k+1} \right) (T_{\circ} - \lambda I) = I$.

3.3 Asymptotic Analysis

In this section we introduce the basic definitions of asymptotic analysis: the concepts of orders of magnitude, of asymptotic approximations and asymptotic expansions. These concepts, although elementary, must be

used carefully, especially when one wishes to study (as we do in subsequent chapters) vector-valued functions $\phi(t, \epsilon)$, $t \in D \subset \mathbb{C}$, $\epsilon \in (0, \epsilon_0]$ for $\epsilon \downarrow 0$.

3.3.1. Orders of magnitude

Consider two real continuous functions $f(\epsilon)$ and $g(\epsilon)$, $\epsilon \in [0, \epsilon_0]$. The behavior of these functions as $\epsilon \downarrow 0$ can be compared using the order symbols O , o and O_s defined below.

Definition 3.3.1

- i) $f=O(g)$ for $\epsilon \downarrow 0$ if there exists positive constants k and C such that $|f(\epsilon)| \leq k|g(\epsilon)|$ for $0 < \epsilon < C$.
- ii) $f=o(g)$ if $\lim_{\epsilon \downarrow 0} f(\epsilon)/g(\epsilon) = 0$.
- iii) $f=O_s(g)$ if $f=O(g)$ and $f \neq o(g)$.

□

We will say that $\delta(\epsilon)$ is an *order function* if it is real, positive, continuous and monotonic, and if $\lim_{\epsilon \downarrow 0} \delta(\epsilon) = 0$. Order functions can be used to analyze the way a function $f(\epsilon)$ tends to zero as $\epsilon \downarrow 0$. We shall almost exclusively use $\delta_n(\epsilon) = \epsilon^n$ as order functions in the chapters that follows.

When studying the asymptotic behavior of vector-valued functions $\phi(t, \epsilon)$ of more than one variable, Definition 3.1 can be applied when considering any arbitrary but fixed value $t = t_0$. In this way we obtain *pointwise* order of magnitude estimates.

Definition 3.3.2 Let $\delta(\epsilon)$ be an order function.

- i) $\phi(t, \epsilon) = O(\delta)$ at $t=t_0$ if there exist positive constants k and C such that $|\phi(t_0, \epsilon)| \leq k\delta(\epsilon)$ for $0 < \epsilon < C$.
- ii) $\phi(t, \epsilon) = o(\delta)$ at $t=t_0$ if $\lim_{\epsilon \downarrow 0} |\phi(t_0, \epsilon)| / \delta(\epsilon) = 0$
- iii) $\phi(t, \epsilon) = O_s(\delta)$ at $t=t_0$ if, at $t=t_0$
 $\phi(t, \epsilon) = O(\delta)$ and $\phi(t, \epsilon) \neq o(\delta)$.

When not only the behavior of $\phi(t, \epsilon)$ at $t=t_0$ for $\epsilon \downarrow 0$ is of interest but also the behavior of $\phi(t, \epsilon)$ in a domain $D \subset \mathbb{C}$, *uniform* order of magnitude estimates are used.

Definition 3.3.3. Let $\delta(\epsilon)$ be an order function.

- i) $\phi(t, \epsilon) = O(\delta)$ *uniformly in D* if there exist constants k and C independent of t such that for all $t \in D$, $|\phi(t, \epsilon)| \leq k \delta(\epsilon)$ for $0 < \epsilon < C$.
- ii) $\phi(t, \epsilon) = o(\delta)$ *uniformly in D* if $\lim_{\epsilon \downarrow 0} |\phi(t, \epsilon)| / \delta(\epsilon) = 0$ uniformly in D .
- iii) $\phi(t, \epsilon) = O_s(\delta)$ *uniformly in D* if $\phi(t, \epsilon) = O(\delta)$ uniformly in D and $\phi(t, \epsilon) \neq o(\delta)$ uniformly in D .

3.3.2 Asymptotic expansions

The objective of asymptotic analysis is to find "good" approximations of a function $\phi(t, \varepsilon)$ for small ε . In this section we make this notion precise.

Definition 3.3.4 i) Let $\phi(t, \varepsilon)$ be a function such that

$\phi(t, \varepsilon) = O_{\varepsilon}(1)$ at $t=t_0$. A function $\phi^a(t, \varepsilon)$ is an

asymptotic approximation of $\phi(t, \varepsilon)$ at $t=t_0$ if

$$\phi(t, \varepsilon) - \phi^a(t, \varepsilon) = o(1) \text{ at } t=t_0.$$

ii) Let $\phi(t, \varepsilon)$ be a function such that $\phi(t, \varepsilon) = O_{\varepsilon}(1)$

uniformly in D . $\phi^a(t, \varepsilon)$ is a *uniform asymptotic*

approximation of $\phi(t, \varepsilon)$ in D if $\phi(t, \varepsilon) - \phi^a(t, \varepsilon) = o(1)$

uniformly in D . □

Let $\{\delta_n(\varepsilon)\}_{n=0}^{\infty}$ be a sequence of order functions such that

$$\delta_{n+1}(\varepsilon) = o(\delta_n(\varepsilon)).$$

Definition 3.3.5 i) $\sum_{n=0}^N \delta_n(\varepsilon) \phi^{(n)}(t, \varepsilon)$ is an *asymptotic*

expansion to $N+1$ terms of $\phi(t, \varepsilon)$ at $t=t_0$ if

$$\phi(t, \varepsilon) - \sum_{n=0}^N \delta_n(\varepsilon) \phi^{(n)}(t, \varepsilon) = o(\delta_{N+1}) \text{ at } t=t_0.$$

ii) $\sum_{n=0}^N \delta_n(\varepsilon) \phi^{(n)}(t, \varepsilon)$ is a *uniform asymptotic expansion*

to $N+1$ terms of $\phi(t, \varepsilon)$ in D if

$$\phi(t, \varepsilon) - \sum_{n=0}^N \delta_n(\varepsilon) \phi^{(n)}(t, \varepsilon) = o(\delta_{N+1}) \text{ uniformly in } D. \quad \square$$

If in the definition above, N can be taken arbitrarily large then infinite asymptotic expansions are obtained. We will denote then by:

$$\phi(t, \epsilon) = \sum_{n=0}^{\infty} \delta_n(\epsilon) \phi^{(n)}(t, \epsilon)$$

although nothing is implied about the convergence of these series.

The question of convergence is of no particular interest in asymptotic theory. It is important, however, to keep in mind the difference between

convergent and asymptotic series: If $\sum_{n=0}^{\infty} \delta_n(\epsilon) \phi^{(n)}(t, \epsilon)$ is a con-

vergent series then for a *given value of* ϵ the difference between

$\sum_{n=0}^N \delta_n(\epsilon) \phi^{(n)}(t, \epsilon)$ and the series limit is arbitrarily small provided

N is large enough; if on the other hand $\sum_{n=0}^{\infty} \delta_n(\epsilon) \phi^{(n)}(t, \epsilon)$ is a series

asymptotic to $\phi(t, \epsilon)$, then *for a given* N , the difference between

$\sum_{n=0}^N \delta_n(\epsilon) \phi^{(n)}(t, \epsilon)$ and $\phi(t, \epsilon)$ is arbitrarily small provided ϵ is small

enough.

3.4 Perturbation Theory in Finite Dimensional Spaces

3.4.1 The problem

Let T be a linear operator on a finite dimensional space V to itself. In this section we consider how the eigenvalues and eigenvectors (or eigenspaces) change when T is subjected to a small perturbation. In

dealing with such a problem it is often convenient to consider a family of operators of the form $T(\epsilon) = T + \epsilon T'$ where ϵ is a scalar parameter supposed to be small. $T(0)=T$ is called the *unperturbed operator* and $\epsilon T'$ the *perturbation*. An interesting question is whether the eigenvalues and eigenvectors of $T(\epsilon)$ can be expressed as a power series in ϵ in the neighborhood of $\epsilon=0$. If this is the case, the change of the eigenvalues and eigenvectors is of the same order of magnitude as the perturbation. As we shall see below, this is not always the case.

In general we will assume that an operator-valued function $T(\epsilon)$ is given which is continuous in a neighborhood of $\epsilon=0$, say for $\epsilon \in [0, \epsilon']$, and we will distinguish two cases:

- a) $T(\epsilon)$ has an *absolutely convergent* power series expansion for $\epsilon \in (0, \epsilon']$, i.e.,

$$T(\epsilon) = T + \sum_{n=1}^{\infty} \epsilon^n T^{(n)} \quad \text{for } \epsilon \in [0, \epsilon'] \quad (3.4.1)$$

$$\|T(\epsilon)\| \leq \|T\| + \sum_{n=1}^{\infty} \epsilon^n \|T^{(n)}\| \triangleq \gamma(\epsilon) < \infty \quad \text{for } \epsilon \in [0, \epsilon'] \quad (3.4.2)$$

- b) $T(\epsilon)$ has an *asymptotic expansion* in powers of ϵ for $\epsilon \rightarrow 0$, i.e.,

$$\lim_{\epsilon \rightarrow 0} \left\| T(\epsilon) - T - \sum_{n=1}^N \epsilon^n T^{(n)} \right\| \cdot \epsilon^{-N} = 0 \quad \forall N \geq 0 \quad (3.4.3)$$

which we will also denote by

$$T(\epsilon) = T + \sum_{n=1}^{\infty} \epsilon^n T^{(n)} \quad (3.4.4)$$

The proof of perturbation results for cases a) and b) are very similar and we will carry them in parallel highlighting the differences whenever they occur. If no mention is made about the convergent or asymptotic character of a series, it will be understood that the result applies to both cases. Obviously, any result proved for b) also applies to a) which will be referred to as the analytic case.

We reproduce here several results in [Kat 66] extended in some cases to the non-analytic (i.e., asymptotic) case.

3.4.2 Perturbation of the eigenvalues

The eigenvalues of $T(\epsilon)$ satisfy the characteristic equation:

$$\det(T(\epsilon) - \xi I) = 0 \quad (3.4.5)$$

If $T(\epsilon)$ is analytic this is an algebraic equation in ξ of degree $n = \dim V$, with coefficients which are analytic in ϵ . It follows from a well known result in function theory (see [Kno 45]) that the roots of (4.5) are branches of analytic functions of ϵ with only algebraic singularities and therefore the number of (distinct) eigenvalues of $T(\epsilon)$ is a constant s independent of ϵ , except at some special values of ϵ . There are only a finite number of such exceptional points in a compact

interval $\epsilon \in [0, \epsilon']$. We will assume that ϵ' is small enough so that $[0, \epsilon']$ contains only one exceptional point which, without loss of generality, we take as $\epsilon=0$.

In a neighborhood of the exceptional point, the eigenvalues of $T(\epsilon)$ can be expressed by s analytic functions $\lambda_1(\epsilon), \dots, \lambda_s(\epsilon)$ with $\lambda_h(\epsilon) \neq \lambda_k(\epsilon)$ for $h \neq k$ which can be grouped in the manner:

$$\{\lambda_1(\epsilon), \dots, \lambda_p(\epsilon)\}, \{\lambda_{p+1}(\epsilon), \dots, \lambda_{p+q}(\epsilon)\}, \dots \quad (3.4.6)$$

in such a way that we have Puiseux series:

$$\lambda_h(\epsilon) = \lambda + \alpha_1 \omega^h \epsilon^{1/p} + \alpha_2 \omega^{2h} \epsilon^{2/p} + \dots \quad (3.4.7)$$

$h=0, 1, \dots, p-1$

where λ is an eigenvalue of the unperturbed operator $T(0)$ and $\omega = \exp[2\pi i/p]$. Each group is called a *cycle* and the number of elements its *period*. It should be noticed that the $\lambda_n(\epsilon)$ are continuous at $\epsilon=0$; $\lambda = \lambda_h(0)$ will be called the *center* of the cycle under consideration.

In general there are several cycles with the same center λ . All eigenvalues belonging to cycles with center λ are said to depart from the unperturbed eigenvalue λ by *splitting* at $\epsilon=0$. The set of these eigenvalues will be called the λ -*group* since they cluster around λ for ϵ small. There is always splitting at and only at an exceptional point (i.e. at $\epsilon=0$ under our assumptions). Equation (4.7) shows that $\lambda_h(\epsilon) - \lambda = O_s(\epsilon^{1/p})$. If $p \geq 2$, therefore, the rate of change at $\epsilon=0$ of the eigenvalues of a cycle of period p is infinitely large as compared with the change in $T(\epsilon)$.

In the non-analytic case the number of eigenvalues may change with ε quite irregularly; the splitting and coalescence of eigenvalues taking place in a very complicated manner. It may even happen that in no interval of the form $(0, \varepsilon']$ is the number of distinct eigenvalues constant. In dealing with non-analytic perturbations we will restrict ourselves to the case of constant number of eigenvalues for $\varepsilon \in (0, \varepsilon']$. This assumption simplifies the analysis and nevertheless provides a sufficiently general setting for applications.

3.4.3 Perturbation of the resolvent

The resolvent of $T(\varepsilon)$

$$R(\xi, T(\varepsilon)) = (T(\varepsilon) - \xi I)^{-1} \tag{3.4.8}$$

is defined for all $\xi \in \mathbb{C}$ not equal to any of the eigenvalues of $T(\varepsilon)$.

The following lemma gives an expression for $R(\xi, T(\varepsilon))$ as a power series in ε with coefficients depending on ξ .

Lemma 3.4.1 Let

$$T(\varepsilon) = T + \sum_{n=1}^{\infty} \varepsilon^n T^{(n)} \tag{3.4.9}$$

If $\xi \in p(T)$ then for, ε small enough, say $\varepsilon \in]0, \varepsilon_0[$,

$\xi \in p(T(\varepsilon))$ and

$$R(\xi, T(\varepsilon)) = R(\xi, T) + \sum_{n=1}^{\infty} \varepsilon^n R^{(n)}(\xi) \tag{3.4.10}$$

where

$$R^{(n)}(\xi) = \sum_{\substack{v_1 + \dots + v_p = n \\ v_i \geq 1}} (-1)^p R(\xi, T)^T \binom{v_1}{R(\xi, T)^T} \binom{v_2}{\dots T} \dots \binom{v_p}{R(\xi, T)^T} \quad (3.4.11)$$

the sum being taken for all combinations of positive integers p and v_1, \dots, v_p such that $1 \leq p \leq n$, $v_1 + \dots + v_p = n$.

The series (4.10) is uniformly convergent on compact subsets of $\rho(T)$ if (4.9) is convergent and it is uniform asymptotic series for $R(\xi, T(\epsilon))$ in compact subsets of $\rho(T(\epsilon))$ if (4.9) is an asymptotic series. □

Proof: Let $\xi \in \rho(T)$ and define $A(\epsilon) = T(\epsilon) - T$. We then have

$$T(\epsilon) - \xi I = (I + A(\epsilon)R(\xi, T))(T - \xi I)$$

Let ϵ'_0 be such that $\|A(\epsilon)R(\xi, T)\| < 1$ for $\epsilon \in [0, \epsilon'_0]$. Then

$$(I + A(\epsilon)R(\xi, T))^{-1} = \sum_{n=0}^{\infty} (-A(\epsilon)R(\xi, T))^n$$

is well defined and we have

$$(T(\epsilon) - \xi I)^{-1} = R(\xi, T)(I + A(\epsilon)R(\xi, T))^{-1}$$

for $\epsilon \in [0, \epsilon'_0]$.

Let now Γ be a compact subset of $\rho(T)$. On Γ , $\|R(\xi, T)\|$ is bounded above and therefore $\|R(\xi, T)\|^{-1}$ has a positive minimum on Γ attained say at $\xi = \xi_0$.

Let ε_0 be such that

$$||T(\varepsilon) - T|| < ||R(\xi_0, T)|| \leq ||R(\xi, T)||^{-1}, \quad \xi \in \Gamma, \quad \varepsilon \in [0, \varepsilon_0]$$

Then $\forall \xi \in \Gamma$ and $\varepsilon \in [0, \varepsilon_0]$,

$$R(\xi, T(\varepsilon)) = R(\xi, T) \sum_{n=0}^{\infty} (-A(\varepsilon)R(\xi, T))^n$$

or

$$R(\xi, T(\varepsilon)) = R(\xi, T) \sum_{n=0}^{\infty} \left[- \left(\sum_{k=1}^N \varepsilon^{k_T(k)} + r_n(\varepsilon) \right) R(\xi, T) \right]^n$$

where $r_N(\varepsilon) \rightarrow 0$ as $N \rightarrow \infty$ for a fixed ε if (4.9) is a convergent series and $r_N(\varepsilon) = o(\varepsilon^N)$ if (4.9) is an asymptotic series. Collecting terms

in like powers of ε we have

$$R(\xi, T(\varepsilon)) = R(\xi, T) + \sum_{n=1}^N \varepsilon^n R^{(n)}(\xi) + r_N(\varepsilon, \xi)$$

where the residual $r_N(\varepsilon, \xi)$ satisfies $||r_N(\varepsilon, \xi)|| \leq ||r_N(\varepsilon, \xi_0)|| \stackrel{\Delta}{=}$

$\bar{r}_N(\varepsilon)$ and $\bar{r}_N(\varepsilon) \rightarrow 0$ as $N \rightarrow \infty$ for a fixed $\varepsilon \in [0, \varepsilon_0]$ in the analytic case and $r_N(\varepsilon)/\varepsilon^N \rightarrow 0$ as $\varepsilon \downarrow 0$ for a fixed N in the non-analytic case.

□

3.4.4 Perturbation of the eigenprojections

Let λ be an eigenvalue of $T=T(0)$, with multiplicity m (meaning algebraic multiplicity unless otherwise stated). Let Γ be a closed positive contour in $\rho(T)$ enclosing λ but no other eigenvalues of T . It follows from Lemma 4.1 that for ε small enough $R(\xi, T(\varepsilon))$ exists for $\xi \in \Gamma$ and therefore there are no eigenvalues of $T(\varepsilon)$ on Γ .

The operator

$$P(\varepsilon) = \frac{1}{2\pi i} \int_{\Gamma} R(\xi, T(\varepsilon)) d\xi \quad (3.4.12)$$

is a projection that commutes with $T(\varepsilon)$ and is equal to the sum of the eigenprojections for all the eigenvalues of $T(\varepsilon)$ lying inside Γ .

Integrating (3.4.10) term by term we get

$$P(\varepsilon) = P + \sum_{n=1}^{\infty} \varepsilon^n P^{(n)} \quad \varepsilon \in [0, \varepsilon_0] \quad (3.4.13)$$

where

$$P = -\frac{1}{2\pi i} \int_{\Gamma} R(\xi, T) d\xi \quad (3.4.14)$$

is the eigenprojection for the eigenvalue λ of T , and

$$P^{(n)} = -\frac{1}{2\pi i} \int_{\Gamma} R^{(n)}(\xi) d\xi \quad (3.4.15)$$

Again, (3.4.13) is convergent in the analytic case and asymptotic in the non-analytic case. In both cases, however, $P(\varepsilon)$ is continuous in a neighborhood of zero and it follows from Lemma 3.2.3 that the range of $P(\varepsilon)$ is isomorphic to the range of P ; in particular,

$$\dim \mathcal{R}(P(\varepsilon)) = \dim \mathcal{R}(P) = m \quad (3.4.16)$$

and since (3.4.16) is true for all ε small enough, it follows that all eigenvalues of $T(\varepsilon)$ lying inside Γ form exactly the λ -group. For this

reason $P(\epsilon)$ will be called the *total projection*, and $R(P(\epsilon))$ the *eigenspace* for the λ -group.

The next lemma gives a series expression for $(T(\epsilon) - \lambda I)P(\epsilon)$ that will play a central role in latter chapters.

Lemma 3.4.2 Let

$$T(\epsilon) = T + \sum_{n=1}^{\infty} \epsilon^n T^{(n)} \quad (3.4.17)$$

Let λ be an eigenvalue of T with multiplicity m and let $P(\epsilon)$ denote the total projection for the λ -group. Then

$$\begin{aligned} \frac{(T(\epsilon) - \lambda)P(\epsilon)}{\epsilon} &= -\frac{1}{2\pi i \epsilon} \int_{\Gamma} (\xi - \lambda) R(\xi, T(\epsilon)) d\xi = \\ &= \frac{D}{\epsilon} + \sum_{n=0}^{\infty} \epsilon^n \tilde{T}^{(n)} \quad \epsilon \in (0, \epsilon_0] \end{aligned} \quad (3.4.18)$$

where Γ is a closed positive contour enclosing λ but no other eigenvalues of T , D is the eigennilpotent for λ and $\tilde{T}^{(n)}$ is given by:

$$\begin{aligned} \tilde{T}^{(n)} &= - \sum_{p=1}^{n+1} (-1)^p \sum_{\substack{v_1 + \dots + v_p = n+1 \\ k_1 + \dots + k_{p+1} = p-1 \\ v_i \geq 1, k_j \geq -m+1}} \binom{k_1}{T} \binom{v_1}{S} \binom{k_2}{S} \dots \binom{k_p}{T} \binom{v_p}{S} \binom{k_{p+1}}{S} \end{aligned} \quad (3.4.19)$$

with $S^{(0)} = -P(0) = -P$, $S^{(-k)} = -D^k$, $k \geq 1$ and $S^{(k)} = S^k$,
 $k \geq 1$ for

$$S = \frac{1}{2\pi i} \int_{\Gamma} (\xi - \lambda)^{-1} R(\xi, T) d\xi \quad (3.4.20)$$

Again, (4.18) is a convergent or an asymptotic series according with the nature of the series (3.4.17).

Proof: Let $\gamma \cup \Gamma$ be a contour that encloses all eigenvalues of T with γ and Γ disjoint and Γ enclosing λ but no other eigenvalue of T . For ϵ small enough $\gamma \cup \Gamma$ encloses all eigenvalues of $T(\epsilon)$ and it follows from (3.2.32) that

$$T(\epsilon) = - \frac{1}{2\pi i} \int_{\gamma \cup \Gamma} \xi R(\xi, T(\epsilon)) d\xi \quad (3.4.21)$$

on the other hand

$$\begin{aligned} T(\epsilon)P(\epsilon) &= \left(\frac{1}{2\pi i}\right)^2 \int_{\Gamma'} \int_{\Gamma} \xi R(\xi, T(\epsilon)) R(\zeta, T(\epsilon)) d\zeta d\xi + \\ &\quad \left(\frac{1}{2\pi i}\right)^2 \int_{\Gamma} \int_{\gamma} \xi R(\xi, T(\epsilon)) R(\zeta, T(\epsilon)) d\zeta d\xi \end{aligned} \quad (3.4.22)$$

where Γ has been slightly expanded to Γ' in the definition of $P(\epsilon)$ to facilitate the integration. Using the resolvent equation (3.2.16) and the Cauchy's integral theorem (3.4.22) becomes

$$T(\epsilon)P(\epsilon) = - \frac{1}{2\pi i} \int_{\Gamma} \xi R(\xi, T(\epsilon)) d\xi \quad (3.4.23)$$

and therefore

$$(T(\epsilon)-\lambda)P(\epsilon) = -\frac{1}{2\pi i} \int_{\Gamma} (\xi-\lambda)R(\xi, T(\epsilon))d\xi \quad (3.4.24)$$

or, using Lemma 3.4.1,

$$\frac{(T(\epsilon)-\lambda)P(\epsilon)}{\epsilon} - \frac{D}{\epsilon} + \sum_{n=0}^N \epsilon^n \tilde{T}^{(n)} = -\frac{1}{2\pi i} \int_{\Gamma} (\xi-\lambda)r_N(\epsilon, \xi)d\xi \quad (3.4.25)$$

where $r_N(\epsilon, \xi)$ is the residual of a convergent or asymptotic series (uniform in any case in Γ) depending on the nature of (3.4.17), and

$$\tilde{T}^{(n)} = -\frac{1}{2\pi i} \sum_{\substack{v_1+\dots+v_p=n+1 \\ v_i \geq 1}} (-1)^p \int_{\Gamma} (\xi-\lambda)R(\xi, T) T^{(v_1)} \dots T^{(v_p)} R(\xi, T) d\xi \quad (3.4.26)$$

To evaluate this integral, we substitute $R(\xi, T)$ by its Laurent expansion (3.2.17) at λ , which we write for convenience in the form:

$$R(\xi, T) = \sum_{k=-m}^{\infty} (\xi-\lambda)^k S^{(k+1)} \quad (3.4.27)$$

with $S^{(0)} = -P$, $S^{(-k)} = D^k$, $k \geq 1$ and $S^{(k)} = S^k$ for S as in (3.4.20).

Substitution of (3.4.27) into the integrand of (3.4.26) gives a Laurent series in $(\xi-\lambda)$ of which only the term with the power $(\xi-\lambda)^{-1}$ contributes to the integral. The result is given by the finite sum (3.4.19). The convergence or asymptotic properties of (3.4.25) follow from the limiting properties of the residual $r_N(\epsilon, \xi)$ and the compactness of Γ . □

In the following chapters we will be mainly interest in (3.4.18) for the case of $\lambda=0$ and semisimple. Under these conditions, it follows from Section 3.2.5 that (3.4.18) reduces to:

$$\frac{T(\epsilon)P(\epsilon)}{\epsilon} = \sum_{n=0}^{\infty} \epsilon^n T^{-1}(n) \quad (3.4.28)$$

for

$$T^{-1}(n) = - \sum_{p=1}^{n+1} (-1)^p \sum_{\substack{v_1 + \dots + v_p = n+1 \\ k_1 + \dots + k_{p+1} = p-1 \\ v_i \geq 1, k_j \geq 0}} S^{(k_1)} T^{(v_1)} \dots T^{(v_p)} S^{(v_{p+1})} \quad (3.4.29)$$

with $S^{(0)} = -P$, $S^{(k)} = (T^\#)^k$. P being the projection on $N(T)$ along $R(T)$ and $T^\#$ the generalized inverse of T .

3.5. Linear Dynamical Systems

We are interested here in linear ordinary differential equations of the form:

$$\begin{cases} \dot{x}(t) = Ax(t) & t \in \mathbb{R}^+ \\ x(0) = x_0 \end{cases} \quad (3.5.1)$$

where $x(t) \in \mathbb{R}^n$ and A is a given linear operator (constant real $(n \times n)$ matrix). The literature of linear dynamical systems is extensive and we do not review it here in detail. Only a few non-standard topics are developed. In Section 3.5.1 we discuss the stability properties, i.e.,

the behavior of $x(t)$ as $t \rightarrow \infty$ of such systems introducing the concept of semistability. The splitting of (3.5.1) into set of lower-dimensional evolutions in invariant subspaces is reviewed in Section 3.5.2. Finally, in Section 3.5.3 the properties of systems which maintain non-negativity of every component of the state vector are analyzed.

3.5.1 Stability and Semistability

We will say that (3.5.1) is a *stable* system if $\|x(t)\| \leq K, t \geq 0$.

If, in addition,

$$\lim_{t \rightarrow \infty} x(t) = 0 \tag{3.5.2}$$

then we will say that the system is *asymptotically stable*. It is well known that (3.5.1) is asymptotically stable if and only if all eigenvalues of A have negative real parts. We will say in this case that A itself is asymptotically stable.

As a generalization of the asymptotic stability concept, consider the following:

Definition 5.1 The system (3.5.1) will be called *semistable* if the following limit exists:

$$x_{\infty} = \lim_{t \rightarrow \infty} x(t) \tag{3.5.3}$$

for any initial state x_0 .

□

The following lemma gives necessary and sufficient conditions on the system matrix A for semistability.

Lemma 3.5.2 i) The system (3.5.1) is semistable iff A has SSNS and all its non-zero eigenvalues have negative real parts. If (3.5.1) is semistable then $\lim_{t \rightarrow \infty} x(t) = P_0 x_0$ where P_0 is the projection on $N(A)$ along $R(A)$.

Proof: i) The solution to (5.1) is given by

$$x(t) = e^{At} x_0$$

Let $A = \sum_{k=0}^s (\lambda_k P_k + D_k)$ be the spectral decomposition of A. Suppose that

$\lambda_0 = 0$ and denote by m_k the multiplicity of λ_k . It follows from (2.34) and (2.35) that

$$x(t) = (P_0 + \tilde{D}_0) x_0 + \sum_{k=1}^s (e^{\lambda_k t} P_k + \tilde{D}_k) x_0$$

where

$$\tilde{D}_k = \sum_{j=1}^{m_k-1} \frac{t^j e^{\lambda_k t}}{j!} (D_k)^j$$

Clearly, if $D_0 = 0$ and $\text{Re } \lambda_k < 0$, $k=1, \dots, s$ then

$$\lim_{t \rightarrow \infty} x(t) = P_0 x_0$$

otherwise the limit does not exist. It follows from Lemma 3.2.5 that P_0 is the projection on $N(A)$ along $R(A)$. □

If a matrix A has SSNS and all its non-zero eigenvalues have negative real parts, A will be called *semistable*.

Lemma 3.5.3

If A is semistable then there exist positive constant α and M such that:

$$\|x(t) - P_0 x_0\| \leq M e^{-\alpha t} \quad (3.5.3)$$

Proof: Follows immediately from the spectral decomposition used in the proof of Lemma 3.5.2. □

Notice that the only difference between stable and semistable systems is the presence of purely imaginary eigenvalues. Thus, if a system is stable and has no purely imaginary eigenvalues then it is semistable.

3.5.2 Flow-invariant subspaces and splitting of evolutions

Consider the linear dynamical system (3.5.1) and let $W \subset \mathbb{R}^n$ be a subspace. We will say that W is a *flow-invariant subspace* of (3.5.1) if $x_0 \in W$ implies $x(t) \in W \quad t \geq 0$. If

$$\mathbb{R}^n = W_1 \oplus \dots \oplus W_r \quad (3.5.4)$$

with W_j flow-invariant subspaces of (3.5.1) we will say that the decomposition (3.5.4) *splits the evolution* (3.5.1).

and therefore the evolution of $x(t)$ can be computed by combining a set of lower dimensional evolutions each taking place in a different flow-invariant subspaces.

3.5.3 Positive Linear Systems

A positive linear system is a linear dynamical system which preserves non-negativity of the state variables, i.e., if $x_0 \geq 0^*$ then $x(t) \geq 0$, $t > 0$. Such systems arise frequently in applications, since in many real systems the state variables represent physical quantities which have no meaning unless they are non-negative. In addition to this practical relevance, positive systems have particular characteristics that makes them specially amenable to analysis.

The following lemma gives necessary and sufficient conditions on the system matrix A for a linear dynamical system to be positive.

Lemma 3.5.1

The system

$$\dot{x}(t) = Ax(t) \quad t > 0$$

is positive if and only if the elements of A , a_{ij} , satisfy:

$$a_{ij} \geq 0 \quad \text{for } i \neq j \quad \square$$

Proof: See [Lue 79]

A matrix will be called *positive* (non-negative) if all its elements are

*Positivity or non-negativity of all elements of a vector or matrix will be denoted by the corresponding sign symbol applied to the vector or to the matrix itself.

positive (non-negative). A matrix with all off-diagonal terms non-negative will be called a *Metzler matrix*. Notice that A is a Metzler matrix iff $A+cI \geq 0$ for some $c > 0$. The theory of positive systems is built upon the rich theory of positive matrices whose cornerstone is the Perron-Frobenius theorem (for several different proofs see [Bel 60]):

Theorem 3.5.2

- i) If $A > 0$ then there exist $\lambda_0 > 0$ and a vector $x_0 > 0$ such that
 - a) $Ax_0 = \lambda_0 x_0$.
 - b) If $\lambda \neq \lambda_0$ is any other eigenvalue of A , then $|\lambda| < \lambda_0$.
 - c) λ_0 is semisimple.
- ii) If $A \geq 0$ then there exist $\lambda_0 \geq 0$ and a vector $x_0 \geq 0$
 - a) $Ax_0 = \lambda_0 x_0$.
 - b) If $\lambda \neq \lambda_0$ is any other eigenvalue of A , then $|\lambda| \leq \lambda_0$.

The equivalent result for Metzler matrices is the following:

Theorem 3.5.3

Let A be a Metzler matrix. Then there exists a real μ_0 and a vector $x_0 \geq 0$ such that:

- a) $Ax_0 = \mu_0 x_0$.
- b) If $\mu \neq \mu_0$ is any other eigenvalue of A , then $\text{Re}(\mu) < \mu_0$.

Proof: See [Lue 79].

□

The Perron-Frobenius theorem thus guarantees the existence of a dominant eigenvalue in positive systems. This eigenvalue and its associated eigenspace determine the long term behavior of a positive system. The following lemma states a stability property of positive systems that we will use in subsequent chapters.

Lemma 3.5.4.

A stable positive system is semistable.

Proof: If the system

$$\dot{x}(t) = Ax(t)$$

is stable then all eigenvalues, λ , of A satisfy $\text{Re}(\lambda) \leq 0$ and if zero is an eigenvalue of A it must be semisimple. If in addition A is a Metzler matrix then it follows from Theorem 5.3 that all non-zero eigenvalues satisfy $\text{Re}(\lambda) < 0$.

□

This corollary indicates that a continuous time stable positive system cannot have sustained oscillations. Notice that even though positivity is apparently a base-dependent characteristic, the fact that a system is positive in some basis has some base-independent implications such as the eigenvalue structure on the imaginary axis.

In the following chapters we will need to deal with systems which preserve non-negativity only for certain initial conditions. Let the decomposition $R^n = W + V$ split a given dynamical system with system matrix A and denote by P the projection on W along V . We will say that the system is *positive in W* if there exists a basis in which the part of $x(t)$ that evolves in W , i.e., $Px(t)$, remains non-negative for any non-negative initial condition. Notice that such a system is positive if the non-negative initial condition is restricted to W . The following lemma gives necessary and sufficient conditions for a dynamical system to be positive in an invariant subspace.

Lemma 3.5.5

Let $R^n = W_1 \oplus \dots \oplus W_r$ be a decomposition in terms of invariant subspaces of A and let $\left\{ P_k \right\}_{k=1}^r$ be the corresponding projections.

i) The system

$$\dot{x}(t) = Ax(t) \quad t \geq 0$$

is positive in W_k if and only if there exists a basis for which $P_k \geq 0$ and $AP_k + cP_k \geq 0$ for some $c > 0$.

Proof: i) Suppose first that $P_k x(t) \geq 0 \quad t \geq 0$ and for arbitrary $x(0) \geq 0$.

For $t > 0$ we have:

$$\frac{1}{t} P_k x(t) = \frac{1}{t} P_k e^{At} x(0) = (AP_k + \frac{1}{t} P_k) x(0) + \frac{o(t)}{t}$$

and therefore for t small enough, say $t < t_0$,

$$AP_k + \frac{1}{t_0} P_k \geq 0$$

Conversely, suppose that $P_k \geq 0$ and $AP_k + cP_k \geq 0$ for some $c > 0$, then:

$$P_k e^{At} = P_k e^{-ct} e^{(A+cI)t} = e^{-ct} P_k \sum_{n=0}^{\infty} \frac{(AP_k + cP_k)^n}{n!} t^n \geq 0$$

□

We next extend Lemma 3.5.4 to systems which are positive in an invariant subspace.

Lemma 3.5.6

Let $\Pi \geq 0$ be a projection commuting with A . If $\Pi e^{At} \geq 0$ and $\|\Pi e^{At}\| \leq K$ then $\Pi A = A\Pi$ is semistable.

Proof: $\Pi e^{At} = e^{A\Pi t} - I + \Pi$ and therefore $\|e^{A\Pi t}\| \leq K$. thus implying that $A\Pi$ is stable. By the positivity condition $\Pi A + c\Pi \geq 0$ (for some $c > 0$) it follows from Theorem 3.5.3 that $\Pi A + c\Pi$ has a real maximal eigenvalue λ_0 or, because $(\Pi A + c\Pi) x_0 = \lambda_0 x_0$ implies $\Pi A x_0 = (\lambda_0 - c)x_0$, that ΠA has a real maximal eigenvalue. Thus, if ΠA is stable it is also semistable. □

This semistability property of stable positive system will play a basic role in Chapter V where it is used to prove aggregation results for such systems.

CHAPTER IV: MULTIPLE TIME SCALE BEHAVIOR OF
SINGULARLY PERTURBED LTI SYSTEMS

4.1 Introduction and Overview

In this chapter we develop a methodology for the asymptotic analysis of linear systems with multiple time scales, establishing the basic results that will be used in subsequent chapters.

In section 4.2 we discuss the relationship among the concepts of weak couplings, singular perturbations and multiple time scale behavior in LTI systems. The focus of our analysis is the vector differential equation:

$$\frac{dx^\varepsilon(t)}{dt} = A(\varepsilon)x^\varepsilon(t) \quad (4.1.1)$$

where ε is a small parameter modeling weak couplings among different parts of the system. We argue that this formulation is appropriate to model systems with phenomena occurring at different time scales, and that it includes most of the formulations used in the past for this purpose. A literature survey is presented in Section 4.3.

Our main results are developed in Sections 4.4. and 4.5 where it is proved that, under a certain "multiple semistability condition" on the system matrix $A(\varepsilon)$, an asymptotic approximation to $x^\varepsilon(t)$, uniformly valid for $t \geq 0$, can be constructed which clearly displays the multiple time scale behavior of $x^\varepsilon(t)$. In Section 4.5 the equations that describe the evolution of $x^\varepsilon(t)$ at different time scales are interpreted as reduced-order models of the system (4.1.1) each valid at the corresponding time scale.

The conditions required for these results to hold are discussed in Section 4.6 where it is argued that they are in fact necessary if $x^\epsilon(t)$ is to have well defined behavior at all time scales. We give examples of systems which do not satisfy these conditions and we suggest extensions of our results to these cases,

4.2 Singularly Perturbed LTI Systems

4.2.1 Regular and singular perturbations, and time scales

We consider here semistable LTI systems of the form:

$$\frac{dx^\epsilon(t)}{dt} = A(\epsilon)x^\epsilon(t), \quad x^\epsilon(0) = x_0 \quad (4.2.1)$$

where $\epsilon \in [0, \epsilon_0]$ $x^\epsilon(t) \in \mathbb{R}^n$ and the matrix $A(\epsilon)$ is assumed to have a power series expansion in ϵ , i.e.,

$$A(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p A_p \quad (4.2.2)$$

This series can be either a convergent series (and we will refer to this as the analytic case), or an asymptotic series. If (4.2.2) is an asymptotic series, then we will also assume that $\text{rank } A(\epsilon)$ is constant for $\epsilon \in (0, \epsilon_0]$ (which it necessarily is in the analytic case). In both cases we will refer to this constant as the *normal rank* of $A(\epsilon)$ and we will denote by $\text{nrnk } A(\epsilon)$.

Our objective is to analyze the behavior of $x^\epsilon(t)$ as $\epsilon \downarrow 0$ on the time interval $[0, \infty)$. The system (4.2.1) can be viewed as a perturbation of

$$\frac{dx^0(t)}{dt} = A_0 x^0(t) \quad x^0(0) = x_0 \quad (4.2.3)$$

and one of the obvious questions to be addressed is that of the relationship between $x^\epsilon(t)$ and $x^0(t)$ for small ϵ . Specifically, under what conditions is $x^0(t)$ a good approximation of $x^\epsilon(t)$? If it is not, how can we construct such an approximation? The first question is resolved in this section and the rest of the chapter deals with the second question.

The following proposition states that $x^0(t)$ is a uniform asymptotic approximation of $x^\epsilon(t)$ on any compact time interval $[0, T]$. (Notice that for this result no stability condition is required).

Proposition 4.2.1

If $A(\epsilon)$ is as in (4.2.2), then

$$\lim_{\epsilon \downarrow 0} \sup_{t \in [0, T]} \| \exp\{A(\epsilon)t\} - \exp\{A_0 t\} \| = 0 \quad (4.2.4)$$

for any $T < \infty$.

Proof: It follows from (3.2.33) and the fact that the eigenvalues of $A(\epsilon)$ are continuous in ϵ for $\epsilon \in [0, \epsilon_0]$, that:

$$\exp\{A(\epsilon)t\} = - \frac{1}{2\pi i} \int_{\Gamma} e^{\lambda t} R(\lambda, A(\epsilon)) d\lambda$$

where Γ is positive contour enclosing all eigenvalues of $A(0)$. From Lemma 3.4.1 it follows that $\|R(\lambda, A(\epsilon)) - R(\lambda, A_0)\| \rightarrow 0$ as $\epsilon \downarrow 0$ uniformly on Γ , and therefore

$$\|\exp\{A(\epsilon)t\} - \exp\{A_0 t\}\| \leq K(\epsilon)e^{\alpha t}$$

for some $K(\epsilon) \rightarrow 0$ as $\epsilon \downarrow 0$ and for some real α , thus proving (4.2.4). \square

In general, however, as the simple example below shows, it is not true that

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \|\exp\{A(\epsilon)t\} - \exp\{A_0 t\}\| = 0 \tag{4.2.5}$$

and therefore, in general, $x^0(t)$ is not a good approximation to $x^\epsilon(t)$ over the infinite interval $[0, \infty)$ no matter how small ϵ is.

Example 4.2.2

Let $A(\epsilon) = -\epsilon$

$$\sup_{t \geq 0} \|\exp\{A(\epsilon)t\} - \exp\{A(0)t\}\| = \sup_{t \geq 0} 1 - e^{-\epsilon t} = 1 \quad \square$$

If eq. (4.2.5) is satisfied, we will say that (4.2.1) is a *regularly perturbed* version of (4.2.3), otherwise we will say that it is *singularly perturbed*. In what follows we will deal primarily with singularly perturbed systems because, as we will now see, it is only in this case that we can talk about different behavior at different time scales. Let us first formalize the notion of multiple time scale behavior.

Definition 4.2.3

Let $x^\epsilon(t)$ be the solution of (4.2.1) and let $\alpha(\epsilon)$ be an order function. We will say that $x^\epsilon(t)$ has a *well defined behavior at time scale $t/\alpha(\epsilon)$* if there exists a bounded continuous function $y(t)$, called the evolution of $x^\epsilon(t)$ at this time scale, such that:

$$\lim_{\epsilon \downarrow 0} \sup_{t \in [\delta, T]} \|x^\epsilon(t/\alpha(\epsilon)) - y(t)\| = 0 \tag{4.2.6}$$

$$\forall \delta > 0, \forall T < \infty, \forall x_0$$

Equivalently, we will say that the LTI system (4.2.1) has a well defined behavior at time scale $t/\alpha(\epsilon)$ if there exists a bounded continuous matrix $Y(t)$ such that

$$\lim_{\epsilon \downarrow 0} \sup_{t \in [\delta, T]} \|\exp\{A(\epsilon)t/\alpha(\epsilon)\} - Y(t)\| = 0 \tag{4.2.7}$$

$$\forall \delta > 0, \forall T < \infty \quad \square$$

According to this definition, the system in Example 4.2.2 has well defined behavior at time scale t/ϵ and its evolution at this time scale is given by:

$$y(t) = e^{-t}$$

Although in this example the convergence condition (4.2.6) is satisfied even for $\delta=0$, it will become clear later on that in general an arbitrarily small interval around zero must be excluded to obtain uniform convergence.

The next proposition shows that semistable regularly perturbed systems have trivial and uninteresting time scale behavior.

Proposition 4.2.4

If $A(\epsilon)$ is a regularly perturbed matrix which is semistable for $\epsilon \in [0, \epsilon_0]$ then, for any order function $\alpha(\epsilon)$,

$$\lim_{\epsilon \rightarrow 0} \sup_{t \in [\delta, T]} \left\| \exp\{A(\epsilon)t/\alpha(\epsilon)\} - P_0 \right\| = 0 \quad (4.2.8)$$

$\forall \delta > 0, \forall T < \infty$

where P_0 is the eigenprojection for the zero eigenvalue of $A_0 = A(0)$.

Proof:

Notice first that:

$$\left\| \exp\{A(\epsilon)t/\alpha(\epsilon)\} - P_0 \right\| \leq \left\| \exp\{A(\epsilon)t/\alpha(\epsilon)\} - \exp\{A_0 t/\alpha(\epsilon)\} \right\| + \left\| \exp\{A_0 t/\alpha(\epsilon)\} - P_0 \right\|$$

It follows from the fact that $A(\epsilon)$ is regularly perturbed that the first term of the right hand side sum converges to zero uniformly for $t \geq 0$ as $\epsilon \rightarrow 0$. To estimate the second term notice that

$$\exp\{A_0 t\} = P_0 - \frac{1}{2\pi i} \int_{\Gamma_0} e^{\lambda t} R(\lambda, A_0) d\lambda$$

where Γ_0 is a positive contour enclosing all non-zero eigenvalues of A_0 . By the semistability property of A_0 , Γ_0 can be chosen to lie

in the left half plane and therefore:

$$\| \exp\{A_0 t/\alpha(\epsilon)\} - P_0 \| \leq K e^{-\beta \cdot \delta/\alpha(\epsilon)} \quad t \in [\delta, \infty]$$

for some $\beta < 0$, thus proving (4.2.8). □

It follows from the above proposition that, if properly modeled, a system with a non-trivial multiple time scale behavior will result in a singularly perturbed o.d.e. The next proposition gives necessary and sufficient conditions on $A(\epsilon)$ for (4.2.1) to be singularly perturbed.

Proposition 4.2.5

The equation (4.2.1), with $A(\epsilon)$ semistable for $\epsilon \in [0, \epsilon_0]$ and of the form (4.2.2), is a singularly perturbed o.d.e. iff rank $A(\epsilon)$ is discontinuous at $\epsilon=0$.

Proof:

Suppose that rank $A(\epsilon) = \text{rank } A_0$ for $\epsilon \in [0, \epsilon_0]$. Then there is no splitting of the zero eigenvalue and therefore, for ϵ small enough, the only singularity of $R(\lambda, A(\epsilon))$, inside a positive contour γ_0 enclosing the origin but no other eigenvalues of A_0 , is at zero. By semisimplicity of the zero eigenvalue of $A(\epsilon)$, this singularity is a simple pole with residual $P_0(\epsilon)$, the eigenprojection for the zero eigenvalue of $A(\epsilon)$.

We thus have:

$$- \frac{1}{2\pi i} \int_{\gamma_0} e^{\lambda t} R(\lambda, A(\epsilon)) d\lambda = P_0(\epsilon) \tag{4.2.9}$$

It follows from eq.(3.4.13) that $P_\circ(\epsilon) \rightarrow P_\circ$ as $\epsilon \downarrow 0$, where P_\circ , the eigenprojection for the zero eigenvalue of A_\circ satisfies:

$$P_\circ = -\frac{1}{2\pi} \int_{\Gamma_\circ} R(\lambda, A_\circ) d\lambda = -\frac{1}{2\pi} \int_{\Gamma_\circ} e^{\lambda t} R(\lambda, A_\circ) d\lambda \quad (4.2.10)$$

We thus have that, for ϵ small enough,

$$\begin{aligned} \|\exp\{A(\epsilon)t\} - \exp\{A_\circ t\}\| &\leq \\ &\frac{1}{2\pi} \int_{\Gamma_\circ} \|R(\lambda, A(\epsilon)) - R(\lambda, A_\circ)\| e^{\operatorname{Re}\lambda t} d\lambda + o(1) \end{aligned}$$

where Γ_\circ is a positive contour enclosing all the non-zero eigenvalues of A_\circ . By the uniform convergence of $R(\lambda, A(\epsilon))$ to $R(\lambda, A_\circ)$ on Γ_\circ and by the semistability of A_\circ that allows us to choose Γ_\circ in the left-half plane, it follows that

$$\lim_{\epsilon \downarrow 0} \sup_{t > 0} \|\exp\{A(\epsilon)t\} - \exp\{A_\circ t\}\| = 0$$

and therefore if (4.2.1) is singularly perturbed then rank $A(\epsilon)$ must be discontinuous at zero. Conversely, suppose that (4.2.1) is regularly perturbed then,

$$\lim_{\epsilon \downarrow 0} P(\epsilon) \stackrel{\Delta}{=} \lim_{\epsilon \downarrow 0} \lim_{t \rightarrow \infty} \exp\{A(\epsilon)t\} = \lim_{t \rightarrow \infty} \exp\{A_\circ t\} = P_\circ \quad (4.2.11)$$

where $P(\epsilon)$ and P_\circ are the eigenprojection for the zero eigenvalue of $A(\epsilon)$ and A_\circ respectively. Thus, rank $P(\epsilon)$ is continuous at $\epsilon=0$. Also, rank $P(\epsilon) = \text{null } A(\epsilon)$ and therefore (4.2.11) implies $\text{null } A(\epsilon) = \text{rank } P(\epsilon)$
rank $P_\circ = \text{null } A_\circ$. □

As a corollary to the above proposition notice that if A_0 is asymptotically stable then *any* perturbation is regular. In effect, if A_0 is asymptotically stable so is $A(\epsilon)$ for ϵ small enough and therefore $\text{null } A_0 = \text{null } A(\epsilon) = 0$ for $\epsilon \in [0, \epsilon_0]$.

We have so far established that the analysis of semistable LTI systems with multiple time scale behavior corresponds to the study of singularly perturbed o.d.e.'s and that the presence of weak couplings may produce well defined behavior at several time scales only if the perturbation changes the rank of the system matrix $A(\epsilon)$. To keep this discussion clear we have only considered systems for which $A(\epsilon)$ is semistable for $\epsilon \in [0, \epsilon_0]$. As we will see in the following sections, this is a necessary (although not sufficient) condition for the system to have well defined behavior at all time scales. Extensions of the results derived for semistable systems to some classes of non-semistable systems are considered in Chapter VI.

The problems on which we focus in the following sections can be formulated as follows:

- i) Under what conditions does a singular perturbed system have well defined behavior at several time scales?
- ii) What are those time scales? Is there a finite number of fundamental time scales at which the system has a non-trivial evolution?
- iii) Is it possible to combine the evolutions of a system at its fundamental time scales to produce a uniform asymptotic approximation to $x^\epsilon(t)$ valid on $[0, \infty)$?

As a way of getting some insight into these questions we discuss in the next section the relationship between the multiple time scale behavior of (4.2.1) and the asymptotic behavior of the eigenvalues of $A(\epsilon)$.

4.2.2 Eigenvalues and time scales

As we mentioned in Chapter I, the notion of multiple time scale behavior of a system is associated with the existence of widely separated eigenvalues. In this section we make the connection between this notion and our analysis of multiple time scale behavior as a problem in singular perturbation.

We again refer to the system

$$\frac{dx^\epsilon(t)}{dt} = A(\epsilon)x^\epsilon(t) , \quad x^\epsilon(0) = x_0 \quad (4.2.12)$$

which is assumed to be semistable for $\epsilon \in [0, \epsilon_0]$ and with system matrix of the form (4.2.2). Assume for simplicity that $A(\epsilon)$ is diagonalizable and let $\lambda_h(\epsilon)$, $h=0,1,\dots,s$ be its distinct eigenvalues. Using the spectral representation of $\exp\{A(\epsilon)t\}$ we get:

$$\exp\{A(\epsilon)t\} = \sum_{h=0}^s e^{\lambda_h(\epsilon)t} P_h(\epsilon) \quad (4.2.13)$$

where $P_h(\epsilon)$ is the eigenprojection for the eigenvalue $\lambda_h(\epsilon)$ of $A(\epsilon)$.

It is clear from (4.2.13) and the semistability assumption that for $\exp\{A(\epsilon)t\}$ to have a non-trivial, well defined behavior at time scale,

say, $t/\alpha(\epsilon)$, it is necessary that there exists some eigenvalues $\lambda_h(\epsilon)$ such that $\lambda_h(\epsilon)/\alpha(\epsilon) \rightarrow \mu_h$ as $\epsilon \downarrow 0$. Thus, for a system to have multiple time scale behavior in the sense of Definition 4.2.3, it must have eigenvalues of different orders of magnitude in ϵ . Our formulation is therefore in accordance with the notion of time scales as a manifestation of eigenvalue separation.

This point of view provides some insight into our discussion in Section 4.2.1 and into the questions raised there. First, the existence of eigenvalues of $A(\epsilon)$ that converge to zero as $\epsilon \downarrow 0$ implies that zero itself must be an eigenvalue of A_0 and that $\text{rank } A(\epsilon)$ must be discontinuous at zero as stated in Proposition 4.2.5. Second, as indicated in (3.4.7), the eigenvalues of $A(\epsilon)$ always have a power series expansion in fractional power of ϵ . Therefore, it is logical to conclude that the fundamental time scales of (4.2.12) must be of the form t/ϵ^q , for some rational q . Furthermore, only a finite number of them can exist (at most one per each distinct eigenvalue).

In addition to the eigenvalue structure, the existence of the limit of $\exp\{A(\epsilon)t/\alpha(\epsilon)\}$ as $\epsilon \downarrow 0$ clearly depends on the structure of the eigenspaces, i.e., on the behavior as $\epsilon \downarrow 0$ of $P_h(\epsilon)$ in (4.2.13). For example, the eigenprojections $P_h(\epsilon)$ (and also the eigennilpotents in the general non-diagonalizable case) have algebraic singularities at $\epsilon=0$ if $\lambda_h(\epsilon) \rightarrow 0$ (see [Kat 66]) and therefore the above limit may not exist

even if there are eigenvalues of order $O_s(\alpha(\epsilon))$. It is this aspect of the time scale problem that is overlooked in the heuristic view of time scales as eigenvalue separation and onto which we will focus our attention in the following sections.

We will first give conditions under which all eigenvalues of $A(\epsilon)$ that converge to zero satisfy

$$\lim_{\epsilon \downarrow 0} \frac{\lambda_h(\epsilon)}{k_h} = \mu_k \quad (4.2.14)$$

for some integer $k_h > 0$, and we will show that systems which satisfy these conditions have well defined behavior at all time scales, with t/ϵ^{k_h} being the fundamental ones. We will also show that if these conditions are violated then at certain time scales $t/\alpha(\epsilon)$ the limit of $\exp\{A(\epsilon)t/\alpha(\epsilon)\}$ as $\epsilon \downarrow 0$ does not exist and we will provide several examples to illustrate what happens in these cases.

4.3 Survey of Related Literature

We review in this section some work of Hoppensteadt and Campbell and Rose ([Hop 75], [Cam 78a,78b,79]), with whom we share a similar problem formulation and similar goals. These treatments have already been briefly discussed in Chapter II; here we will focus on the connections between their results and ours with specific references to later results in this chapter.

In [Hop 75] the author deals with singularly perturbed evolutions in Banach spaces of the form:

$$\dot{v}(t) = A(\epsilon)v(t) + G(v(t), \epsilon) \quad (4.3.1)$$

For the purposes of comparison with our work, we will take $G(v, \epsilon) \equiv 0$ and will assume $A(\epsilon)$ to be a finite dimensional operator. Hoppensteadt starts by assuming that $A(\epsilon)$ is stable for $\epsilon > 0$ and that it can be written in the form

$$A(\epsilon) = A_0(\epsilon) \oplus \epsilon^{r_1} A_1(\epsilon) \oplus \dots \oplus \epsilon^{r_m} A_m(\epsilon) \quad (4.3.2)$$

with $A_i(\epsilon)$ also stable and continuous in ϵ . He then constructs a *uniform asymptotic approximation* of the solution of

$$v^\epsilon(t) = A(\epsilon)v^\epsilon(t) \quad (4.3.3)$$

of the form:

$$\tilde{v}^\epsilon(t) = \sum_{i=0}^m V_i(\sigma_i) + o(1) \quad (4.3.4)$$

where $\sigma_j = \epsilon^{r_j} t$ are the different time scales and $V_j(t)$ are the solutions of:

$$\dot{V}_j(t) = A_j(0)V_j(t) \quad j=0, \dots, m \quad (4.3.5)$$

It is thus established that (4.3.3) has well defined behavior at time scales t/ϵ^{r_j} $j=1, \dots, m$ in the sense of Definition 4.2.3, and that the behavior of (4.3.3) at different time scales needs to be considered

in a uniform asymptotic approximation. In fact, Hoppensteadt is the *only* author known to us to have considered asymptotic approximations uniformly valid on $[0, \infty)$.

The crucial assumption made by Hoppensteadt in equation (4.3.2) is avoided in our work. We start with an *arbitrary* matrix $A(\epsilon)$ and we give a *necessary and sufficient condition* for a uniform asymptotic approximation of $\exp\{A(\epsilon)t\}$ in terms of several time scales to exist. We give a *constructive* method to compute the decomposition (4.3.2) and the integers r_j , $j=1, \dots, m$ that determine the different time scales. See Section 4.4.1 and more specifically equations (4.4.26)-(4.4.28). From an applications viewpoint this represents a major advantage since our starting point is closer to the way real systems are specified.

Campbell and Rose in [Cam 79] analyze the limiting behavior of:

$$\exp\{(A+\epsilon B)t/\epsilon\} \quad (4.3.6)$$

as $\epsilon \downarrow 0$ proving that a well defined pointwise limit exists if and only if A is semistable. In this case the limit is given by:

$$\lim_{\epsilon \downarrow 0} e^{(A/\epsilon+B)t} = e^{(I-AA^D)Bt} (I-AA^D) \quad (4.3.7)$$

where A^D is the Drazin inverse of A . This result gives a necessary and sufficient condition for $A(\epsilon) = A+\epsilon B$ to have well-defined behavior at time scale t/ϵ . Given the pointwise nature of the limit and the fact that attention is focused a priori on one particular time scale,

nothing is implied about the existence of well defined behavior at slower time scales. One major consequence of this is that the question of a uniform asymptotic approximation of $\exp\{A(\epsilon)t\}$ is not addressed.

In [Cam 78] the case of

$$A(\epsilon) = A + \epsilon^r B + \epsilon^{s-1} C + \epsilon^s D \quad (4.3.8)$$

for $s > r > 1$ is analyzed focusing now on time scale t/ϵ^s . The objective is to determine conditions under which the pointwise limit

$$\lim_{\epsilon \downarrow 0} \exp\{A(\epsilon)t/\epsilon^s\} \quad (4.3.9)$$

exists, i.e., conditions are sought under which $A(\epsilon)$ will have well defined behavior at time scale t/ϵ^s . Several sufficient conditions for (4.3.9) to exist are obtained in terms of the matrices A , B and C but the development is quite complicated and no underlying pattern is readily apparent or exposed when there is behavior at several time scales.

Campbell, in summary, concentrates his attention in obtaining *pointwise limits for a given time scale*. The question of how many and which time scales a system exhibits and the construction of uniform asymptotic approximations obtained by combining the different pointwise limits is not addressed in his work. The limiting results obtained by Campbell are also obtained as corollaries of Theorem 4.4.4 proved in the next section (in particular, see Section 4.4.3). We also show

that the conditions under which we prove well-defined multiple time scale behavior are necessary and sufficient. (see Section 4.4.4).

This chapter can thus be seen as a clarification and significant extension of the work of Hoppensteadt and Campbell in three ways:

- 1) In *deriving* rather than *assuming* a decomposition of $A(\epsilon)$ that clearly displays its time scale structure. In this way we overcome the criticism often directed at work such as that of Hoppensteadt that assumes an *a-priori* knowledge of the time scale structure of the system.
- 2) In establishing a *necessary and sufficient* condition for the existence of well defined multiple time scale behavior.
- 3) In producing a *uniform* asymptotic approximation rather than pointwise approximations.

4.4 The Multiple Semistability Case: Complete Time Scale Decomposition

In this section we study singularly perturbed LTI systems^(*)

$$\frac{dx^\epsilon(t)}{dt} = A_0(\epsilon)x^\epsilon(t) \quad (4.4.1)$$

which satisfy a multiple semistability condition to be specified in Section 4.4.1. These systems have well defined behavior at all time scales and have non-trivial evolution at a finite number of fundamental time scales which are of the form t/ϵ^k , for a set of integers k . In Sections 4.4.2 and 4.4.3 we compute explicitly the evolutions of (4.4.1) at the fundamental time scales and we show that they can be used

- a) to construct an asymptotic approximation of $x^\epsilon(t)$ uniformly valid on $[0, \infty)$ and
- b) to define a set of reduced-order models of (4.4.1) each valid at a different time scale.

Finally, in Section 4.4 we show that such a complete time scale decomposition is possible only under the multiple semistability condition.

4.4.1 The multiple semisimple null structure and the multiple semistability conditions

Let $A_0(\epsilon)$, $\epsilon \in [0, \epsilon_0]$ be a semistable matrix with a series expansion (convergent or asymptotic) of the form:

$$A_0(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p A_{0p} \quad (4.4.2)$$

(*) We change here our notation from $A(\epsilon)$ to $A_0(\epsilon)$ because we need to construct a sequence of matrices $A_k(\epsilon)$ the first of which will be the system matrix in (4.1).

In this section we specify two conditions that will be seen to play a basic role in the and normal rank d multiple time scale behavior of (4.4.2). For our development we first need to construct a sequence of matrices $A_k(\epsilon)$, $k=1, \dots, m$, obtained recursively from $A_0(\epsilon)$ as indicated below.

Let $P_0(\epsilon)$ denote the total projection for the zero-group of eigenvalues of $A_0(\epsilon)$. It follows from Lemma 3.4.2 that, if A_{00} has SSNS, the matrix

$$A_1(\epsilon) \triangleq \frac{P_0(\epsilon)A_0(\epsilon)}{\epsilon} = \frac{A_0(\epsilon)P_0(\epsilon)}{\epsilon} = \frac{P_0(\epsilon)A_0(\epsilon)P_0(\epsilon)}{\epsilon} \quad (4.4.3)$$

has a series expansion of the form:

$$A_1(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p A_{1p} \quad (4.4.4)$$

If the resulting first order term in (4.4.4), A_{10} , also has SSNS we define the next matrix $A_2(\epsilon)$ as in (4.4.3), i.e.,

$$A_2(\epsilon) \triangleq \frac{P_1(\epsilon)A_1(\epsilon)}{\epsilon} = \frac{P_1(\epsilon)P_0(\epsilon)A_0(\epsilon)}{\epsilon^2} = \sum_{p=0}^{\infty} \epsilon^p A_{2p} \quad (4.4.5)$$

where $P_1(\epsilon)$ is the total projection for the zero-group of eigenvalues of $A_1(\epsilon)$. The recursion ends at

$$\begin{aligned} A_m(\epsilon) &\triangleq \frac{P_{m-1}(\epsilon)A_{m-1}(\epsilon)}{\epsilon} = \frac{P_{m-1}(\epsilon)P_{m-2}(\epsilon)\dots P_0(\epsilon)A_0(\epsilon)}{\epsilon^m} \\ &= \sum_{p=0}^{\infty} \epsilon^p A_{mp} \end{aligned} \quad (4.4.6)$$

if the matrix A_{m0} does not have SSNS.

Before proceeding with several propositions that establish properties of the sequence of matrices $A_k(\epsilon)$ and $P_k(\epsilon)$, we provide some insight into the recursive process outlined above, and a preview of the results that follow. The construction of the matrices $A_k(\epsilon)$ can be interpreted as a way of grouping the eigenvalues of $A_0(\epsilon)$ according to their asymptotic behavior as $\epsilon \downarrow 0$. By definition, $P_0(\epsilon)$ is the sum of the eigenprojections for the eigenvalues of $A_0(\epsilon)$ that converge to zero as $\epsilon \downarrow 0$ (the zero-group). Thus, $Q_0(\epsilon) = I - P_0(\epsilon)$ is the sum of the eigenprojections corresponding to the order one eigenvalues of $A_0(\epsilon)$. The subspaces $R(P_0(\epsilon))$ and $R(Q_0(\epsilon))$ are invariant subspaces of $A_0(\epsilon)$, and $P_0(\epsilon)A_0(\epsilon)$ is the part of $A_0(\epsilon)$ in $R(P_0(\epsilon))$ in the sense that all eigenvalues of $A_0(\epsilon)$ in the zero group are also eigenvalues of $P_0(\epsilon)A_0(\epsilon)$ (with the same eigenspaces) but none of the order one eigenvalues of $A_0(\epsilon)$ is an eigenvalue of $P_0(\epsilon)A_0(\epsilon)$. The matrix $A_1(\epsilon) = P_0(\epsilon)A_0(\epsilon)/\epsilon$ has the following eigenstructure: an eigenvalue at zero with eigenspace $R(Q_0(\epsilon)) \oplus N(A_0(\epsilon))$ and the nonzero eigenvalues that are the zero-group eigenvalues of $A_0(\epsilon)$ scaled by a factor of $1/\epsilon$ with the corresponding eigenspaces unchanged. For example, the order one eigenvalue of $A_1(\epsilon)$ are the order ϵ eigenvalues of $A_0(\epsilon)$ and $R(Q_1(\epsilon))$ is the corresponding eigenspace. At the next step we construct the matrix

$$A_2(\epsilon) = \frac{P_1(\epsilon)A_1(\epsilon)}{\epsilon} = \frac{P_1(\epsilon)P_0(\epsilon)A_0(\epsilon)}{\epsilon^2}$$

which has a zero eigenvalue with eigenspace $R(Q_0(\epsilon)) \oplus R(Q_1(\epsilon))$. Its nonzero eigenvalues are the eigenvalues of $A_0(\epsilon)$ that converge to zero faster than ϵ , scaled by $1/\epsilon^2$. Thus, the order one eigenvalues of $A_2(\epsilon)$ are the order ϵ^2 eigenvalues of $A_2(\epsilon)$ and the corresponding eigenspace in $R(Q_2(\epsilon))$.

This recursive grouping is possible as long as the semisimple null structure condition is satisfied at each stage. If this condition is not violated then for some $m \leq \text{rank } A_0(\epsilon)$, $A_{m+1}(\epsilon) = 0$, i.e., the recursion has a finite number of stages and each eigenvalue of $A_0(\epsilon)$ belongs to one of the above mentioned groups. The case where the recursion ends at $k=m$ because $A_{k,0}$ does not have SSNS is discussed in Section 4.5.3.

The following proposition establishes several basic properties of the matrices $A_k(\epsilon)$ and $P_k(\epsilon)$.

Proposition 4.4.1

Let $A_k(\epsilon)$, $k=0,1,\dots,m$, be the sequence of matrices constructed above, and let $P_k(\epsilon)$, $k=0,1,\dots,m$ be the corresponding total projections for the zero-groups of eigenvalues. Define $Q_k(\epsilon) = I - P_k(\epsilon)$. Then, for ϵ small enough, say $\epsilon \in [0, \epsilon_1]$, we have:

$$i) \quad P_i(\epsilon)P_j(\epsilon) = P_j(\epsilon)P_i(\epsilon) \quad i, j=0, 1, \dots, m \quad (4.4.7)$$

$$ii) \quad Q_i(\epsilon)Q_j(\epsilon)=0 \quad i \neq j \quad i, j=0, 1, \dots, m \quad (4.4.8)$$

$$iii) \quad R^n = R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_k(\epsilon)) \oplus R(P_0(\epsilon) \dots P_k(\epsilon)) \quad (4.4.9)$$

$$k=0, 1, \dots, m$$

$$iv) \quad \text{rank } Q_k(\epsilon) = \text{rank } A_{k,0} \quad (4.4.10)$$

and for $\epsilon \in (0, \epsilon_1]$,

$$v) \quad Q_k(\epsilon)A_0(\epsilon) = Q_k(\epsilon)\epsilon^k A_k(\epsilon) = \epsilon^k A_k(\epsilon)Q_k(\epsilon) \quad (4.4.11)$$

$$= A_0(\epsilon)Q_k(\epsilon) \quad k=0, 1, \dots, m$$

Proof:

Remember that if $\lambda \in \rho(A_{k0})$ then, for ϵ small enough, $\lambda \in \rho(A_k(\epsilon))$.

In what follows, all statements involving ϵ apply, unless otherwise stated, for $\epsilon \in [0, \epsilon_1]$ with ϵ_1 assumed small enough so that $\lambda \in \rho(A_{k0})$ implies $\lambda \in \rho(A_k(\epsilon))$ for $k=0, 1, \dots, m$.

To prove i) notice that by definition of $A_1(\epsilon)$, $P_0(\epsilon)$ and $A_1(\epsilon)$ commute and therefore for $\lambda \in \rho(A_{10})$,

$$(A_1(\epsilon) - \lambda I)P_0(\epsilon)R(\lambda, A_1(\epsilon)) = P_0(\epsilon) =$$

$$(A_1(\epsilon) - \lambda I)R(\lambda, A_1(\epsilon))P_0(\epsilon)$$

or

$$(A_1(\epsilon) - \lambda I)[P_0(\epsilon)R(\lambda, A_1(\epsilon)) - R(\lambda, A_1(\epsilon))P_0(\epsilon)] = 0$$

which implies that $P_0(\epsilon)$ and $R(\lambda, A_1(\epsilon))$ commute because $(A_1(\epsilon) - \lambda I)$ is invertible. We thus have

$$\begin{aligned} P_0(\epsilon)P_1(\epsilon) &= -\frac{1}{2\pi i} \int_{\gamma_1} P_0(\epsilon)R(\lambda, A_1(\epsilon))d\lambda = \\ &= -\frac{1}{2\pi i} \int_{\gamma_1} R(\lambda, A_1(\epsilon))d\lambda P_0(\epsilon) = P_1(\epsilon)P_0(\epsilon) \end{aligned} \quad (4.4.12)$$

where γ_1 is a positive contour enclosing the origin but no eigenvalues of A_{10} different than zero.

To prove ii) notice first two facts:

a) For $\lambda \in \rho(A_{10})$

$$P_0(\epsilon)R(\lambda, A_1(\epsilon)) = R(\lambda, A_1(\epsilon)) + \frac{Q_0(\epsilon)}{\lambda} \quad (4.4.13)$$

which is obtained by multiplying both sides of:

$$P_0(\epsilon) = I - \frac{Q_0(\epsilon)}{\lambda} (\lambda I - A_1(\epsilon))$$

by $R(\lambda, A_1(\epsilon))$. (The above inequality follows from the fact that $Q_0(\epsilon)A_1(\epsilon) = 0$).

b) For $\lambda \in \rho(A_1(\epsilon))$,

$$P_0(\epsilon)R(\lambda, A_1(\epsilon)) = \epsilon R(\epsilon\lambda, A_0(\epsilon))P_0(\epsilon) \quad (4.4.14)$$

which is obtained by multiplying both sides of

$$(A_0(\epsilon) - \epsilon\lambda I)P_0(\epsilon)R(\lambda, A_1(\epsilon)) = \epsilon P_0(\epsilon)$$

by $R(\epsilon\lambda, A_0(\epsilon))$.

Combining (4.4.13) and (4.4.14) we get:

$$R(\lambda, A_1(\epsilon)) = R(\epsilon\lambda, A_0(\epsilon))P_0(\epsilon) - \frac{Q_0(\epsilon)}{\lambda} \quad (4.4.15)$$

Now, it follows from the definition of $Q_1(\epsilon) \triangleq I - P_1(\epsilon)$ and (3.4.12) that

$$Q_1(\epsilon) = - \frac{1}{2\pi i} \int_{\Gamma_1} R(\lambda, A_1(\epsilon)) d\lambda$$

where Γ_1 is a positive contour enclosing all eigenvalues of A_{10} except the zero eigenvalue, and therefore using (4.4.15) we have:

$$Q_1(\epsilon) = - \frac{1}{2\pi i} \int_{\Gamma_1} \epsilon R(\epsilon\lambda, A_0(\epsilon)) P_0(\epsilon) d\lambda + \frac{1}{2\pi i} \int_{\Gamma_1} \frac{Q_0(\lambda)}{\lambda} d\lambda \quad (4.4.16)$$

The second integral is clearly zero because Γ_1 does not enclose the origin, and multiplying (4.4.16) by $Q_0(\epsilon)$ we get:

$$Q_1(\epsilon)Q_0(\epsilon) = - \frac{1}{2\pi i} \int_{\Gamma_1} \epsilon R(\epsilon\lambda, A_0(\epsilon)) d\lambda P_0(\epsilon)Q_0(\epsilon) = 0$$

That $Q_0(\epsilon)$ and $Q_1(\epsilon)$ commute follows from (4.4.12). The direct sum iii) follows immediately from the fact, that $Q_0(\epsilon)$, $Q_1(\epsilon)$ and $P_0(\epsilon)P_1(\epsilon)$ are projections that annihilate each other and add up the identity.

We now proceed with the recursion assuming that i)-iii) are satisfied for $m=l$. From the commutativity property of $P_k(\epsilon)$, $k=0,1,\dots,l$, and the definition of $A_{l+1}(\epsilon)$ it follows that

$$P_j(\epsilon)A_{l+1}(\epsilon) = A_{l+1}(\epsilon)P_j(\epsilon) = A_{l+1}(\epsilon) \quad j \leq l+1 .$$

and therefore for $\lambda \in \rho(A_{\ell+1,0})$ we can write:

$$(A_{\ell+1}(\varepsilon) - \lambda I)P_j(\varepsilon)R(\lambda, A_{\ell+1}(\varepsilon)) = P_j(\varepsilon) = (A_{\ell+1}(\varepsilon) - \lambda I)R(\lambda, A_{\ell+1}(\varepsilon))P_j(\varepsilon)$$

$j \leq \ell+1$

As before, the invertibility of $(A_{\ell+1}(\varepsilon) - \lambda I)$ implies that $P_j(\varepsilon)$ and $R(\lambda, A_{\ell+1}(\varepsilon))$ commute $\forall j \leq \ell+1$ and therefore i) is satisfied for $\ell=m+1$.

To prove ii) for $\ell=m+1$ notice that a recursive use of (4.4.15) gives:

$$R(\lambda, A_{\ell+1}(\varepsilon)) = \varepsilon^{\ell+1} R(\varepsilon^{\ell+1} \lambda, A_0(\varepsilon)) P_0(\varepsilon) P_1(\varepsilon) \dots P_\ell(\varepsilon) -$$

$$\frac{1}{\lambda} [Q_{\ell+1}(\varepsilon) + \varepsilon Q_\ell(\varepsilon) P_{\ell+1}(\varepsilon) + \varepsilon^2 Q_{\ell-1}(\varepsilon) P_\ell(\varepsilon) P_{\ell+1}(\varepsilon) + \dots$$

$$+ \varepsilon^\ell Q_0(\varepsilon) P_1(\varepsilon) \dots P_{\ell+1}(\varepsilon)]$$

We thus have:

$$Q_{\ell+1}(\varepsilon) = - \frac{1}{2\pi i} \int_{\Gamma_{\ell+1}} \varepsilon^{\ell+1} R(\varepsilon^{\ell+1} \lambda, A_0(\varepsilon)) d\lambda P_0(\varepsilon) \dots P_\ell(\varepsilon)$$

readily giving $Q_{\ell+1}(\varepsilon)Q_j(\varepsilon) = 0, \forall j \leq \ell$ and proving that ii) is also satisfied for $m=\ell+1$.

Finally, to extend iii) to $m=\ell+1$ notice that using the fact that

$$P_j(\varepsilon)Q_{\ell+1}(\varepsilon) = Q_{\ell+1}(\varepsilon) \forall j \leq \ell \quad \text{we get}$$

$$Q_{\ell+1}(\varepsilon) + P_0(\varepsilon) \dots P_{\ell+1}(\varepsilon) =$$

$$Q_{\ell+1}(\varepsilon) + P_0(\varepsilon) \dots P_{\ell}(\varepsilon) - P_0(\varepsilon) \dots P_{\ell}(\varepsilon) Q_{\ell+1}(\varepsilon) =$$

$$P_0(\varepsilon) \dots P_{\ell}(\varepsilon)$$

or

$$I = \sum_{j=0}^{\ell} Q_j(\varepsilon) + P_0(\varepsilon) \dots P_{\ell}(\varepsilon) =$$

$$\sum_{j=0}^{\ell+1} Q_j(\varepsilon) + P_0(\varepsilon) \dots P_{\ell+1}(\varepsilon)$$

which together with i) and ii) gives iii).

To prove iv) notice that $Q_k(\varepsilon)$ is continuous at $\varepsilon=0$ and therefore, by Lemma 3.2.4 and for ε small enough, we have:

$$\text{rank } Q_k(\varepsilon) = \text{rank } Q_k(0) = n - \text{rank } P_k = \text{rank } A_{k0}$$

where $P_k = \lim_{\varepsilon \downarrow 0} P_k(\varepsilon)$ is the projection on $N(A_{k0})$ along $R(A_{k0})$.

To prove v) notice that ii) implies:

$$Q_k(\varepsilon) P_j(\varepsilon) = P_j(\varepsilon) Q_k(\varepsilon) = \begin{cases} Q_k(\varepsilon) & k \neq j \\ 0 & k = j \end{cases}$$

and therefore

$$Q_k(\varepsilon) A_0(\varepsilon) = Q_k(\varepsilon) P_{k-1}(\varepsilon) \dots P_0(\varepsilon) A_0(\varepsilon) = Q_k(\varepsilon) \varepsilon^k A_k(\varepsilon)$$

The following proposition establishes that the sequence $A_k(\epsilon)$ always terminates at some finite m .

Proposition 4.4.2

Let $A_k(\epsilon)$, $k=0,1,\dots$ be the sequence of matrices defined by (4.4.6). One of the following two conditions (or both) occur at some finite m :

- i) $A_{m,0}$ does not have SSNS
- ii) $A_{m+1}(\epsilon) = 0^*$

In the analytic case, ii) is equivalent to:

$$\text{ii')} \quad \sum_{k=0}^m \text{rank } A_{k,0} = d \quad (4.4.17)$$

Proof: We only need to prove that ii) or, equivalently, ii') will occur if i) does not. Suppose that all matrices $A_{k,0}$, $k \geq 0$ have SSNS, then for $\forall j \geq 0$,

$$R^n = R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_j(\epsilon)) \oplus R(P_0(\epsilon) \dots P_j(\epsilon)) \quad (4.4.18)$$

and because $\text{rank } Q_k(\epsilon) = \text{rank } A_{k,0}$ only a finite number of $A_{k,0}$'s can be nonzero. Let m be such that $A_{m,0} \neq 0$ and $A_{k,0} = 0$, $k > m$. Notice that $P_k(\epsilon) = I$ if $A_{k,0} = 0$ and therefore $A_{k,0} = 0$ for $k > m$ implies $A_{m+1}(\epsilon) = 0$

($A_{m+1}(\epsilon) \sim 0$ if (4.4.2) is not convergent). Consider now the analytic case.

(*) To be interpreted as $A_{m+1}(\epsilon) \sim 0$ if (4.4.2) is an asymptotic series.

Because

$$A_{m+1}(\epsilon) = \frac{A_0(\epsilon)P_0(\epsilon)P_1(\epsilon)\dots P_m(\epsilon)}{\epsilon},$$

$A_{m+1}(\epsilon)=0$ implies $R(P_0(\epsilon)\dots P_m(\epsilon)) \subset N(A_0(\epsilon))$. On the other hand, if

$x \in N(A_0(\epsilon))$ then $x \in N(A_k(\epsilon))$ and therefore $P_k(\epsilon)x=x$. Consequently,

$N(A_0(\epsilon)) = R(P_0(\epsilon)\dots P_m(\epsilon))$ and from (4.4.18) it follows that

ii) implies ii'). Conversely, if ii') is satisfied, it follows from

(4.4.18) that $\dim R(P_0(\epsilon)\dots P_m(\epsilon)) = n-d = \dim N(A_0(\epsilon))$ which together

with $N(A_0(\epsilon)) \subset R(P_0(\epsilon)\dots P_m(\epsilon))$ again implies $N(A_0(\epsilon)) = R(P_0(\epsilon)\dots P_m(\epsilon))$

and therefore $A_{m+1}(\epsilon)=0$. □

We will say that a matrix $A_0(\epsilon)$ satisfies the *multiple semisimple nullstructure (MSSNS) condition* if the sequence of matrices $A_k(\epsilon)$ can be constructed up to a stage $k=m$ for which (4.4.17) is satisfied with all matrices

$$A_{k,0} = \lim_{\epsilon \rightarrow 0} \frac{P_{k-1}(\epsilon)\dots P_0(\epsilon)A_0(\epsilon)}{\epsilon^k} \quad k=0,1,\dots,m$$

having SSNS. If $A_0(\epsilon)$ satisfies the MSSNS condition and in addition all matrices $A_{k,0}$, $k=0,1,\dots,m$, are semistable we will say that $A_0(\epsilon)$ satisfies *multiple semistability (MSST) condition*. Although in rest of Section 4.4 we mainly deal with matrices which satisfy the MSST condition, all the preliminary results developed in this section hold for the less restrictive

MSSNS condition and will be used in Section 4.5 to partially extend some of the asymptotic results derived for MSST systems to a larger class of systems.

The following proposition provides some insight into the structure of the matrices $A_k(\epsilon)$, $k=1, \dots, m$, and into how they relate to $A_0(\epsilon)$. (See also Figure 4.1 for a graphic diagram of the geometric content of Proposition 4.4.3).

Proposition 4.4.3

If $A_0(\epsilon)$ satisfies the MSSNS condition then, for some $\epsilon_1 > 0$,

i) $A_k(\epsilon)$, $k=0, 1, \dots, m$ have SSNS for $\epsilon \in [0, \epsilon_1]$.

ii) For $\epsilon \in (0, \epsilon_1]$,

$$R(A_k(\epsilon)) = R(Q_k(\epsilon)) \oplus \dots \oplus R(Q_m(\epsilon)) \quad k=0, \dots, m \quad (4.4.19)$$

$$N(A_k(\epsilon)) = R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_{k-1}(\epsilon)) \quad N(A_0(\epsilon)) \quad (4.4.20)$$

$k=1, \dots, m,$

$$N(A_0(\epsilon)) = R(P_0(\epsilon) \dots P_m(\epsilon)) \quad (4.4.21)$$

iii) If $\lambda(\epsilon)$ is an eigenvalue of $A_k(\epsilon)$ not belonging to its zero-group then $\epsilon^k \lambda(\epsilon)$ is an eigenvalue of $A_0(\epsilon)$ in $R(Q_k(\epsilon))$. Conversely, if $\mu(\epsilon)$ is an eigenvalue of $A_0(\epsilon)$ in $R(Q_k(\epsilon))$ then $\epsilon^{-k} \mu(\epsilon)$ is an eigenvalue of $A_k(\epsilon)$ not belonging to its zero-group.

$$\begin{aligned} \mathbb{R}^n &= \mathcal{R}(Q_m(\epsilon)) \oplus \dots \oplus \mathcal{R}(Q_1(\epsilon)) \oplus \mathcal{R}(Q_0(\epsilon)) \oplus \mathcal{N}(A_0(\epsilon)) \\ &\quad \underbrace{\hspace{10em}}_{\mathcal{R}(A_0(\epsilon))} \quad \underbrace{\hspace{10em}}_{\mathcal{N}(A_0(\epsilon))} \\ &\quad \underbrace{\hspace{10em}}_{\mathcal{R}(A_1(\epsilon))} \quad \underbrace{\hspace{10em}}_{\mathcal{N}(A_1(\epsilon))} \\ &\quad \underbrace{\hspace{10em}}_{\mathcal{R}(A_2(\epsilon))} \quad \underbrace{\hspace{10em}}_{\mathcal{N}(A_2(\epsilon))} \\ &\quad \quad \quad \vdots \quad \quad \quad \vdots \\ &\quad \quad \quad \vdots \quad \quad \quad \vdots \\ &\quad \quad \quad \vdots \quad \quad \quad \vdots \\ &\quad \underbrace{\hspace{10em}}_{\mathcal{R}(A_m(\epsilon))} \quad \underbrace{\hspace{10em}}_{\mathcal{N}(A_m(\epsilon))} \end{aligned}$$

Figure 4.1: Illustration of Proposition 4.4.3.

Proof:

We have already seen in the proof of Proposition 4.4.2 that (4.4.17) implies (4.4.21). Now, if $y \in R(A_0(\epsilon))$ then $y = A_0(\epsilon)x$ and it follows from (4.4.9) and the fact that $P_k(\epsilon)A_0(\epsilon) = A_0(\epsilon)P_k(\epsilon)$, that

$$y = \sum_{k=0}^m Q_k(\epsilon)A_0(\epsilon)x$$

which implies $y \in R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_m(\epsilon))$. This together with

$$\text{rank } A_0(\epsilon) = \sum_{k=0}^m \text{rank } A_{k0} = \sum_{k=0}^m \dim R(Q_k(\epsilon))$$

implies

$$R(A_0(\epsilon)) = R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_m(\epsilon)) \quad (4.4.22)$$

which finishes the proof of i), ii) and iii) for $k=0$.

Consider now $N(A_k(\epsilon))$. By definition of $A_k(\epsilon)$, we clearly have that, for ϵ small enough but not equal to zero,

$$N(A_k(\epsilon)) \supset N(A_0(\epsilon)) \oplus R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_{k-1}(\epsilon)) \quad (4.4.23)$$

Establish inclusion in the other direction by contradiction. Let

$x \in N(A_k(\epsilon))$ but not to the right hand side of (4.4.23). From (4.4.9)

we have:

$$R^n = R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_m(\epsilon)) \oplus N(A_0(\epsilon)) \quad (4.4.24)$$

Hence, write $x = x_1 + x_2$ with $x_1 \in R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_{k-1}(\epsilon)) \oplus$

$N(A_0(\epsilon))$ and $0 \neq x_2 \in R(Q_k(\epsilon)) \oplus \dots \oplus R(Q_m(\epsilon))$ with $P_\ell(\epsilon)x = x$ if $\ell < k$.

Now, $x \in N(A_k(\epsilon))$ implies that

$$\begin{aligned} 0 = A_k(\epsilon)x &= \frac{A_0(\epsilon)P_0(\epsilon)\dots P_{k-1}(\epsilon)}{\epsilon^k} x \\ &= \frac{A_0(\epsilon)P_0(\epsilon)\dots P_{k-1}(\epsilon)}{\epsilon^k} x_2 \\ &= \frac{A_0(\epsilon)}{\epsilon^k} x_2 \end{aligned}$$

i.e., $x_2 \in N(A_0(\epsilon))$ thereby yielding a contradiction. This establishes

(4.4.20). To prove (4.4.19), note that by definition of $A_k(\epsilon)$

$$R(A_k(\epsilon)) \subset R(P_0(\epsilon)\dots P_{k-1}(\epsilon)) \cap R(A_0(\epsilon))$$

and, it follows from Proposition 4.4.1 and the SSNS of $A_0(\epsilon)$ that

$$R(P_0(\epsilon)\dots P_{k-1}(\epsilon)) \cap R(A_0(\epsilon)) = R(Q_k(\epsilon)) \oplus \dots \oplus R(Q_m(\epsilon))$$

and

$$\begin{aligned} \dim R(A_k(\epsilon)) &= n - \dim N(A_k(\epsilon)) \\ &= n - \left(\sum_{j=0}^{k-1} \text{rank } Q_j(\epsilon) + (n-d) \right) \\ &= \sum_{j=k}^m \text{rank } Q_j(\epsilon) \\ &= \dim R(P_0(\epsilon)\dots P_{k-1}(\epsilon)) \cap R(A_0(\epsilon)) \end{aligned}$$

This implies that

$$R(A_k(\varepsilon)) = R(Q_k(\varepsilon)) \oplus \dots \oplus R(Q_m(\varepsilon)) \quad (4.4.25)$$

Equation (4.4.24) together with (4.4.20) and (4.4.25) gives

$$R^n = R(A_k(\varepsilon)) \oplus N(A_k(\varepsilon)) \quad k=0,1,\dots,m$$

proving that for ε small enough but not zero the matrices $A_k(\varepsilon)$, $k=0,1,\dots,m$ have SSNS. By the MSSNS condition, the matrices $A_k(0)$ for $k=0,1,\dots,m$ also have SSNS which completes the proof of i). To prove iii) notice that if $A_k(\varepsilon)u = \lambda(\varepsilon)u$ and $\lambda(\varepsilon)$ does not belong to the zero group of eigenvalues of $A_k(\varepsilon)$ then $Q_k(\varepsilon)u = u$ and therefore it follows from that:

$$\begin{aligned} A_0(\varepsilon)u &= A_0(\varepsilon)Q_k(\varepsilon)u = \varepsilon^k A_k(\varepsilon)Q_k(\varepsilon)u = \\ &= \varepsilon^k A_k(\varepsilon)u = \varepsilon^k \lambda(\varepsilon)u \end{aligned}$$

Conversely, let $A_0(\varepsilon)u = \mu(\varepsilon)u$ with $u \in R(Q_k(\varepsilon))$ then,

$$\varepsilon^{-k} \mu(\varepsilon)u = \varepsilon^{-k} A_0(\varepsilon)Q_k(\varepsilon)u = A_k(\varepsilon)Q_k(\varepsilon)u = A_k(\varepsilon)u .$$

It follows from (4.4.11) and (4.4.21) that, if $A_0(\varepsilon)$ has MSSNS then

$$A_0(\varepsilon) = \sum_{k=0}^m \varepsilon^k A_k(\varepsilon)Q_k(\varepsilon) \quad (4.4.26)$$

This decomposition can be interpreted as a decomposition of $A_0(\varepsilon)$

according to the direct sum

$$R^n = R(Q_0(\varepsilon)) \oplus \dots \oplus R(Q_m(\varepsilon)) \oplus R(P_0(\varepsilon) \dots P_m(\varepsilon)) \quad (4.4.27)$$

as follows:

$$A_{\circ}(\epsilon) = \hat{A}_{\circ}(\epsilon) \oplus \epsilon \hat{A}_1(\epsilon) \oplus \dots \oplus \epsilon^m \hat{A}_m(\epsilon) \oplus 0 \quad (4.4.28)$$

where $\hat{A}_k(\epsilon)$ denotes the restriction of $A_k(\epsilon)$ on $R(Q_k(\epsilon))$.

The eigenvalues of $A_{\circ}(\epsilon)$ can thus be divided into $(m+1)$ groups corresponding to the eigenvalues of $A_{\circ}(\epsilon)$ in each of the invariant subspaces $R(Q_k(\epsilon))$. Each eigenvalue of $A_{\circ}(\epsilon)$ is of order ϵ^j for some integer $j \geq 0$ and the eigenvalues of order ϵ^k coincide with ϵ^k times the eigenvalues of order one of $A_k(\epsilon)$. Figure 4.1 illustrates the structure of the matrix $A_k(\epsilon)$: its null space includes, in addition to the null space of $A_{\circ}(\epsilon)$, the eigenspace of $A_{\circ}(\epsilon)$ corresponding to all eigenvalues of order $0(1), 0(\epsilon), \dots, 0(\epsilon^{k-1})$ while its range includes the eigenspaces of $A_{\circ}(\epsilon)$ for all eigenvalues of order $o(\epsilon^{k-1})$. The construction of the sequence $A_k(\epsilon)$ can thus be viewed as a way to separate the eigenvalues of $A_{\circ}(\epsilon)$ in different groups according to their asymptotic order as $\epsilon \downarrow 0$. The actual calculations required to compute the matrices $A_k(\epsilon)$ will be discussed in Section 4.4.5.

The following theorem illustrates the consequences of the MSSNS condition for the multiple time scale behavior of $\exp\{A_{\circ}(\epsilon)t\}$.

Theorem 4.4.4

If $A_{\circ}(\epsilon)$ satisfies the MSSNS condition then:

$$\exp\{A_{\circ}(\epsilon)t\} = \sum_{k=0}^m Q_k(\epsilon) \exp\{A_k(\epsilon)\epsilon^k t\} + P_0(\epsilon) \dots P_m(\epsilon) \quad (4.4.29)$$

$$= \sum_{k=0}^m \exp\{Q_k(\epsilon)A_k(\epsilon)\epsilon^k t\} - mI \quad (4.4.30)$$

$$= \prod_{k=0}^m \exp\{Q_k(\varepsilon)A_k(\varepsilon)\varepsilon^k t\} \quad (4.4.31)$$

$$= \exp\left\{\sum_{k=0}^m Q_k(\varepsilon)A_k(\varepsilon)\varepsilon^k t\right\} \quad (4.4.32)$$

Proof:

Write

$$\exp\{A_0(\varepsilon)t\} = P_0(\varepsilon)\exp\{A_0(\varepsilon)t\} + Q_0(\varepsilon)\exp\{A_0(\varepsilon)t\}$$

A simple algebraic manipulation gives (remember, $P_0(\varepsilon)$ is a projection that commutes with $A_0(\varepsilon)$):

$$\begin{aligned} \exp\{A_0(\varepsilon)t\} &= \exp\{A_1(\varepsilon)\varepsilon t\} - Q_0(\varepsilon) + \\ & \quad Q_0(\varepsilon)\exp\{A_0(\varepsilon)t\} \end{aligned} \quad (4.4.33)$$

Repeating the same manipulation with $\exp\{A_1(\varepsilon)t\}$ we obtain

$$\begin{aligned} \exp\{A_0(\varepsilon)t\} &= \exp\{A_2(\varepsilon)\varepsilon^2 t\} - Q_1(\varepsilon) - Q_0(\varepsilon) + \\ & \quad Q_1(\varepsilon)\exp\{A_1(\varepsilon)\varepsilon t\} + Q_0(\varepsilon)\exp\{A_0(\varepsilon)t\} \end{aligned}$$

By the MSSNS condition satisfied by $A_0(\varepsilon)$ this process can be repeated m times giving:

$$\begin{aligned} \exp\{A_0(\varepsilon)t\} &= \exp\{A_{m+1}(\varepsilon)\varepsilon^{m+1} t\} - \sum_{k=0}^m Q_k(\varepsilon) + \\ & \quad \sum_{k=0}^m Q_k(\varepsilon)\exp\{A_k(\varepsilon)\varepsilon^k t\} \end{aligned} \quad (4.4.34)$$

As we have already shown in the proof of Proposition 4.4.2 the rank condition (4.4.17) implies $A_{m+1}(\varepsilon)=0$ and therefore (4.4.34)

proves (4.4.29). Using the identity

$$Q_k(\epsilon) \exp\{A_k(\epsilon) \epsilon^k t\} = \exp\{Q_k(\epsilon) A_k(\epsilon) \epsilon^k t\} - I + Q_k(\epsilon)$$

in (4.4.29) and the fact that $P_0(\epsilon) \dots P_m(\epsilon) = I - \sum_{k=0}^m Q_k(\epsilon)$ we get (4.4.30).

Equation (4.4.28) gives (4.4.32) and the product formula (4.4.31) follows now from (4.4.32) by the commutativity properties of $Q_k(\epsilon) A_k(\epsilon)$:

$$Q_k(\epsilon) A_k(\epsilon) \cdot Q_j(\epsilon) A_j(\epsilon) = 0 \quad j \neq k$$

which results from (4.4.8) and (4.4.11). □

Equation (4.4.29) corresponds to the splitting of $\exp\{A_0(\epsilon)t\}$ according to the direct sum decomposition (4.4.26). Under the condition of MSSNS, this splitting corresponds also to a decomposition into parts of $\exp\{A_0(\epsilon)t\}$ that evolve at different time scales. For example, $Q_k(\epsilon) \exp\{A_k(\epsilon) \epsilon^k t\}$ does not change significantly until t is of order $1/\epsilon^k$. Theorem 4.4.4 thus gives a consistent spatial and temporal decomposition of eq. (4.4.1) which as we will see is very convenient to study the multiple time scale behavior of (4.4.1) and also to derive uniform asymptotic approximations of $\exp\{A_0(\epsilon)t\}$.

4.4.2 Uniform asymptotic approximation of $\exp\{A_0(\epsilon)t\}$

As we have proved in Proposition 4.2.1, $\exp\{A_0(t)\}$ is a uniform asymptotic approximation to $\exp\{A(\epsilon)t\}$ on any compact time interval $[0, T]$. It is quite clear, however, that this approximation does not capture the

multiple time scale behavior of a singularly perturbed system. To construct an approximation which captures this behavior, we have to require it to be uniformly valid over the infinite time interval $[0, \infty)$.

The next theorem gives the desired approximation under the assumption that $A_0(\epsilon)$ satisfies the MSST condition.

Theorem 4.4.5

Let $A_0(\epsilon)$ satisfy the MSST condition and let $A_k(\epsilon)$, $P_k(\epsilon)$ and $Q_k(\epsilon)$, $k=0,1,\dots,m$, be the sequences of matrices constructed in Section 4.4.1. Then,

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \|\exp\{A_0(\epsilon)t\} - \phi(t, \epsilon)\| = 0 \quad (4.4.35)$$

where $\phi(t, \epsilon)$ is any of the following expressions:

$$\phi(t, \epsilon) = \sum_{k=0}^m Q_k \exp\{A_{k,0} \epsilon^k t\} + P_0 \dots P_m \quad (4.4.36)$$

$$= \sum_{k=0}^m \exp\{A_{k,0} \epsilon^k t\} - mI \quad (4.4.37)$$

$$= \prod_{k=0}^m \exp\{A_{k,0} \epsilon^k t\} \quad (4.4.38)$$

$$= \exp\left\{ \sum_{k=0}^m A_{k,0} \epsilon^k t \right\} \quad (4.4.39)$$

with $A_{k,0} = \lim_{\epsilon \downarrow 0} A_k(\epsilon)$, $P_k = \lim_{\epsilon \downarrow 0} P_k(\epsilon)$ and

$Q_k = \lim_{\epsilon \downarrow 0} Q_k(\epsilon)$.

Furthermore,

$$R^n = R(A_{00}) \oplus \dots \oplus R(A_{m,0}) \oplus \left(\bigcap_{k=0}^m N(A_{k,0}) \right) \quad (4.4.40)$$

Proof:

Let us first prove (4.4.35) for $\phi(t, \epsilon)$ as in (4.4.36).

Using Theorem 4.4.4 we can write:

$$\begin{aligned} \exp\{A_0(\epsilon)t\} - \phi(t, \epsilon) &= (P_0(\epsilon) \dots P_m(\epsilon) - P_0 \dots P_m) + \\ &\sum_{k=0}^m (Q_k(\epsilon) \exp\{A_k(\epsilon)\epsilon^k t\} - Q_k \exp\{A_{k,0}\epsilon^k t\}) \end{aligned}$$

There is no problem with

$$P_0(\epsilon) \dots P_m(\epsilon) - P_0 \dots P_m$$

because $\lim_{\epsilon \downarrow 0} (P_k(\epsilon) - P_k) = 0$ by definition. To estimate each of the other terms write:

$$\begin{aligned} \psi_k(t, \epsilon) &\triangleq Q_k(\epsilon) \exp\{A_k(\epsilon)\epsilon^k t\} - Q_k \exp\{A_{k,0}\epsilon^k t\} = \\ &= -\frac{1}{2\pi i} \int_{\Gamma_k} e^{\lambda \epsilon^k t} (R(\lambda, A_k(\epsilon)) - R(\lambda, A_{k,0})) d\lambda \end{aligned}$$

where Γ_k is a contour that encloses all non-zero eigenvalues of $A_{k,0}$. By semistability of $A_{k,0}$, Γ_k can be taken to lie strictly inside the left half plane and therefore for some $\alpha < 0$

$$||\psi_k(t, \epsilon)|| \leq \frac{1}{2\pi} e^{\alpha \epsilon^k t} \int_{\Gamma_k} ||R(\lambda, A_k(\epsilon)) - R(\lambda, A_{k,0})|| d\lambda$$

which gives

$$||\psi_k(t, \epsilon)|| \leq \frac{1}{2\pi} \int_{\Gamma_k} ||R(\lambda, A_k(\epsilon)) - R(\lambda, A_{k,0})|| d\lambda$$

From Lemma 3.4.1 it follows that $R(\lambda, A_k(\epsilon))$ converges to $R(\lambda, A_{k,0})$ uniformly on Γ_k and therefore

$$||\psi_k(t, \epsilon)|| \leq K(\epsilon) \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0$$

proving (4.4.35). To establish (4.4.40) notice that from (4.4.9) we have

$$R^n = R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_m(\epsilon)) \oplus R(P_0(\epsilon) \dots P_m(\epsilon))$$

and using the fact that $Q_k(\epsilon)$ and $P_k(\epsilon)$ are projections which are continuous in ϵ , we obtain

$$R^n = R(Q_0) \oplus \dots \oplus R(Q_m) \oplus R(P_0 \dots P_m) \tag{4.4.42}$$

By definition Q_k is the projection on $R(A_{k,0})$ along $N(A_{k,0})$ and therefore $R(Q_k) = R(A_{k,0})$ thus (4.4.42) can also be written as

$$R^n = R(A_{0,0}) \oplus \dots \oplus R(A_{m,0}) \oplus \left(\bigcap_{k=0}^m N(A_{k,0}) \right) \tag{4.4.43}$$

Now,

$$Q_k \exp\{A_{k,0} \epsilon^k t\} = \exp\{A_{k,0} \epsilon^k t\} - I + Q_k$$

which proves (4.4.37). The equivalence of the expressions (4.4.37), (4.4.38) and (4.4.39) follows immediately from (4.4.43) i.e., from the fact that

$$A_{k,0} A_{j,0} = 0 \quad j \neq k$$

□

As the above theorem shows, the sequence of matrices $A_{k,0}$, $k=0,1,\dots,m$, completely determines an asymptotic approximation to $\exp\{A_0(\epsilon)t\}$ which captures its multiple time scale behavior. In the next section we use this result to determine the complete multiple time scale behavior of

$$\frac{dx^\epsilon(t)}{dt} = A_0(\epsilon)x^\epsilon(t) \quad (4.4.44)$$

and to define a set of reduced-order models of (4.4.44).

4.4.3 Multiple time scale behavior and reduced-order models

In this section we use Theorem 4.4.5 to show that systems which satisfy the MSST condition have well defined behavior at all time scales and that the matrices $A_{k,0}$ determine a set of reduced-order models of the system.

The following corollary of Theorem 4.4.5 gives an explicit formula for the evolutions of $\exp\{A_0(\epsilon)t\}$.

Corollary 4.4.6

Let $A_o(\epsilon)$ satisfy the MSST condition and let $A_{k,o}, P_k$ and $Q_k, k=0,1,\dots,m$, be the sequence of matrices specified in Theorem 4.4.5. Then,

$$\begin{aligned} \text{i) } \quad \lim_{\epsilon \downarrow 0} \sup_{\delta < t \leq T} \left| \exp\{A_o(\epsilon)t/\epsilon^k\} - \Phi_k(t) \right| &= 0 \\ & \forall \delta > 0, \forall T < \infty \\ & k=0,1,\dots,m-1 \end{aligned} \quad (4.4.45)$$

$$\begin{aligned} \text{ii) } \quad \lim_{\epsilon \downarrow 0} \sup_{\delta < t < \infty} \left| \exp\{A_o(\epsilon)t/\epsilon^m\} - \Phi_m(t) \right| &= 0 \\ & \forall \delta > 0 \end{aligned} \quad (4.4.46)$$

where $\Phi_k(t)$ is either of the following expressions:

$$\Phi_k(t) = Q_k \exp\{A_{k,o}t\} + P_o \dots P_k \quad (4.4.47)$$

$$\begin{aligned} &= P_o \dots P_{k-1} \exp\{A_{k,o}t\} \\ & k=0,1,\dots,m \end{aligned} \quad (4.4.48)$$

Proof

From Theorem 4.4.5 we have

$$\begin{aligned} \exp\{A_o(\epsilon)t/\epsilon^k\} &= \sum_{\ell=0}^{k-1} Q_\ell \exp\{A_{\ell,o}t/\epsilon^{k-\ell}\} + Q_k \exp\{A_{k,o}t\} + \\ &+ \sum_{\ell=k+1}^m Q_\ell \exp\{A_{\ell,o}t/\epsilon^{\ell-k}\} + P_o \dots P_m + o(1) \end{aligned} \quad (4.4.49)$$

uniformly for $t \in [0, \infty)$. Now, by the semistability of $A_{\ell,o}$ we have

$$Q_\ell \exp\{A_{\ell,o}t\} = -\frac{1}{2\pi i} \int_{\Gamma_\ell} e^{\lambda t} R(\lambda, A_{\ell,o}) d\lambda$$

for some Γ_ℓ in the left-half plane, and by boundedness of $R(\lambda, A_{\ell,0})$ on Γ_ℓ ,

$$||Q_\ell \exp\{A_{\ell,0} t\}|| \leq M_\ell e^{-\alpha_\ell t}$$

for some $M_\ell, \alpha_\ell > 0$. We thus get

$$\sup_{\delta < t < \infty} ||Q_\ell \exp\{A_{\ell,0} t/\epsilon^{k-\ell}\}|| \leq M_\ell e^{-\alpha_\ell \delta/\epsilon^{k-\ell}}$$

which gives

$$\lim_{\epsilon \downarrow 0} \sup_{\delta < t < \infty} \sum_{\ell=0}^{k-1} Q_\ell \exp\{A_{\ell,0} t/\epsilon^{k-\ell}\} = 0 \quad (4.4.50)$$

On the other hand, it is clear that

$$\lim_{\epsilon \downarrow 0} \sup_{0 < t \leq T} ||\exp\{A_{\ell,0} \epsilon^{\ell-k} t\} - I|| = 0 \quad \begin{matrix} \forall \ell > k \\ \forall T < \infty \end{matrix} \quad (4.4.51)$$

and using (4.4.50) and (4.4.51) in (4.4.49) we get (4.4.45) and (4.4.46) with

$$\Phi_k(t) = Q_k \exp\{A_{k,0} t\} + \sum_{\ell=k+1}^m Q_\ell + P_0 \dots P_m$$

Equality of this expression with (4.4.47) and (4.4.48) follows from the direct sum decomposition (4.4.42). □

It is now immediate from (4.4.47) and (4.4.36) that the evolutions of $\exp\{A_0(\varepsilon)t\}$ at time scales t/ε^k , $\Phi_k(t)$, $k=0,1,\dots,m$ can be combined to produce a uniform asymptotic approximation to $\exp\{A_0(\varepsilon)t\}$ as follows:

$$\exp\{A_0(\varepsilon)t\} = \sum_{k=0}^m \Phi_k(\varepsilon^k t) - \sum_{k=0}^{m-1} P_0 P_1 \dots P_k + o(1) \quad (4.4.52)$$

This equation shows that only the behavior at time scales t/ε^k , $k=0,1,\dots,m$, is needed to capture the main features of the evolution of $\exp\{A_0(\varepsilon)t\}$ over the infinite time interval $[0,\infty)$. It is clear from the proof of Corollary 4.4.6, however, that the limit

$$\lim_{\varepsilon \downarrow 0} \exp\{A_0(\varepsilon)t/\alpha(\varepsilon)\}$$

exists for any order function $\alpha(\varepsilon)$. Indeed, if $\alpha_k(\varepsilon) = o(\varepsilon^k)$ and $\varepsilon^{k+1} = o(\alpha_k(\varepsilon))$, $k=0,1,\dots,m-1$, then

$$\lim_{\varepsilon \downarrow 0} \exp\{A_0(\varepsilon)t/\alpha_k(\varepsilon)\} = P_0 \dots P_k$$

and for $\alpha(\varepsilon) = o(\varepsilon^m)$,

$$\lim_{\varepsilon \downarrow 0} \exp\{A_0(\varepsilon)t/\alpha(\varepsilon)\} = P_0 \dots P_m$$

Thus the system has well defined behavior at all time scales even though only a finite number of them, that we will call *the fundamental time scales*, are required to capture the main features of the system's evolution.

At a given time scale all modes that evolve at faster time scales will have reached steady state (remember the multiple semistability condition) while modes which evolve at slower time scales have yet to depart from their initial conditions. A reduced-order model derived from $A_{k,0}$ can then be associated with the modes that evolve exactly at time scale t/ϵ^k .

To interpret the matrices $A_{k,0}$ as reduced-order models of the system

$$\frac{dx^\epsilon(t)}{dt} = A_0(\epsilon)x^\epsilon(t) \quad x^\epsilon(0) = x_0 \quad (4.4.53)$$

valid at different time scales, notice the asymptotic approximation

$$\exp\{A_0(\epsilon)t\} = \sum_{k=0}^m Q_k \exp\{A_{k,0}\epsilon^k t\} + P_0 \dots P_m + o(1) \quad (4.4.54)$$

and the direct sum decomposition:

$$R^n = R(Q_0) \oplus \dots \oplus R(Q_m) \oplus R(P_0 \dots P_m)$$

imply that if $x^\epsilon(0) \in R(Q_k)$, then $x^\epsilon(t)$ remains in $R(Q_k)$ for all $t > 0$

except for terms which are uniformly negligible as $\epsilon \downarrow 0$. Thus

$R(Q_k)$, $k=0,1,\dots,m$ and $R(P_0 \dots P_m)$ can be thought of as *almost invariant subspaces* of the system (4.4.53). Furthermore, the parts of $x^\epsilon(t)$ that

evolve in different subspaces do so at different time scales. To describe the part of $x^\epsilon(t)$ that evolves at time scale t/ϵ^k to first order approximation, the following model can be used:

$$\frac{dy_k(t)}{dt} = A_{k,0} y_k(t) \quad k=0,1,\dots,m \quad (4.4.55)$$

If $y_k(0) = Q_k x_0$ then

$$y_k(\epsilon^k t) = Q_k x^\epsilon(t) + o(1) \quad k=0,1,\dots,m \quad (4.4.56)$$

uniformly for $t \geq 0$, and once again a uniform approximation of $x^\epsilon(t)$ can be constructed by combining the solutions of the reduced-order models (4.4.55) as follows:

$$x^\epsilon(t) = \sum_{k=0}^m y_k(\epsilon^k t) + P_0 \dots P_m x_0 + o(1) \quad (4.4.57)$$

Notice also that:

$$\sum_{k=0}^m \text{rank } A_{k,0} = \text{rank } A_0(\epsilon) \quad (4.4.58)$$

and therefore the combined dimensionality of the reduced-order models (4.4.55) equals the dimension of the exact model.

This decomposition of (4.4.53) into a set of reduced-order models is more easily visualized using an appropriate change of basis. From Theorem 4.5.1 we have

$$R^n = R(A_{0,o}) \oplus \dots \oplus R(A_{m,o}) \oplus \left(\sum_{k=0}^m N(A_{k,o}) \right) \quad (4.4.59)$$

and by the SSNS property of the matrices $A_{k,o}$ it follows that:

$$N(A_{k,o}) = R(A_{0,o}) \oplus \dots \oplus R(A_{k-1,o}) \oplus R(A_{k+1,o}) \oplus \dots \oplus R(A_{m,o}) \oplus \left(\sum_{k=0}^m N(A_{k,o}) \right)$$

If we now choose a basis adapted to (4.4.59), the matrix $A_{k,o}$ will have, in this new basis, a block diagonal form with only one non-zero block. That is, if T denotes the change of basis matrix, then

$$TA_{k,o}T^{-1} = \text{diag}\{0,0,\dots,0,\tilde{A}_k,0,\dots,0\}$$

where \tilde{A}_k is a full rank square matrix of dimension equal to $\text{rank } A_{k,o}$.

Using this change of basis we can write (4.4.39) as follows:

$$\begin{aligned} \exp\{A_o(\epsilon)t\} &= T^{-1} \exp \left\{ \sum_{k=0}^m TA_{k,o}T^{-1} \epsilon^k t \right\} T + o(1) \\ &= T^{-1} \text{diag} \left\{ e^{\tilde{A}_0 t}, e^{\tilde{A}_1 \epsilon t}, \dots, e^{\tilde{A}_m \epsilon^m t}, I \right\} T + o(1) \end{aligned} \quad (4.4.60)$$

showing that, to first order approximation, the system (4.4.53) can be thought of composed of $(m+1)$ *uncoupled* subsystems

$$\frac{d\tilde{y}_k(t)}{dt} = \tilde{A}_{k,k} \tilde{y}_k(t) \quad k=0,1,\dots,m$$

each running at a different time scale.

In the next section we show that such a complete decomposition is possible only if $A_0(\epsilon)$ satisfies the MSST condition and in Section 4.4.5 we describe an algorithm for the computation of the matrices $A_{k,0}$.

4.4.4 Necessity of the multiple semistability condition

In Sections 4.4.2 and 4.4.3 we have shown that if a singularly perturbed LTI system satisfies the MSST condition then it has well defined behavior at all time scales and that there exist a set of ϵ -independent reduced-order models which describe the system's evolution at each of its fundamental time scales. Here we will show that if the MSST condition is violated then at least for some time scale, $t/\alpha(\epsilon)$, the limit

$$\lim_{\epsilon \downarrow 0} \exp\{A_0(\epsilon)t/\alpha(\epsilon)\}$$

does not exist. In this case a complete time scale decomposition of the type developed in previous sections is not possible. Some partial extensions to systems that violate the MSST condition are, however, possible and will be considered in Section 4.5 and in Chapter VI.

Two examples will shed some light as to what happens when the MSST condition is violated:

Example 4.4.7

Consider the matrix

$$A_o(\epsilon) = \begin{bmatrix} \epsilon & 0 & -2\epsilon \\ \epsilon & \epsilon & -2\epsilon \\ 1 & 1 & -2 \end{bmatrix}$$

It is semistable for $\epsilon \in [0,1]$ and it has three real eigenvalues

$$\lambda_o = 0; \quad \lambda_1 = -2 + o(1); \quad \lambda_2 = -\epsilon^2 + o(\epsilon^2)$$

The matrices $A_{o,o}$ and $A_{1,o}$ (see Section 4.4.5 for an algorithm to compute them) are given by:

$$A_{o,o} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & -2 \end{bmatrix}; \quad A_{1,o} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & -1/2 & 0 \end{bmatrix}$$

and the MSST condition is violated because $A_{1,o}$ does not have SSNS (it is nilpotent). A direct computation of $\exp\{A_o(\epsilon)t\}$ gives:

$$\exp\{A_o(\epsilon)t\} = \begin{bmatrix} \frac{(\lambda_2 - \lambda_1)}{2} & -\frac{\epsilon}{\lambda_1} (e^{\lambda_1 t} - 1) & \frac{\epsilon(\epsilon - \lambda_1)}{\lambda_1} (e^{\lambda_1 t} - 1) \\ 0 & (\epsilon - \lambda_1 - 2)e^{\lambda_1 t} & 2\epsilon e^{\lambda_1 t} \\ 0 & -e^{\lambda_1 t} & (\epsilon - \lambda_1)e^{\lambda_1 t} \end{bmatrix} \frac{1}{\lambda_2 - \lambda_1} +$$

$$+ \frac{1}{\lambda_2 - \lambda_1} \begin{bmatrix} \frac{\lambda_2 - \lambda_1}{2} & \frac{\epsilon}{\lambda_2} (e^{\lambda_2 t} - 1) & \frac{\epsilon(\lambda_2 - \epsilon)}{\lambda_2} (e^{\lambda_2 t} - 1) \\ 0 & (\lambda_2 - \epsilon + 2)e^{\lambda_2 t} & -2\epsilon e^{\lambda_2 t} \\ 0 & e^{\lambda_2 t} & (\lambda_2 - \epsilon)e^{\lambda_2 t} \end{bmatrix} \quad (4.4.61)$$

and we have the following time scale behavior:

$$\lim_{\epsilon \downarrow 0} \exp\{A_o(\epsilon)t\} = \exp\{A_{o,o}t\} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & (1 - e^{-2t})/2 & e^{-2t} \end{bmatrix}$$

$$\lim_{\epsilon \downarrow 0} \exp\{A_o(\epsilon)t/\epsilon\} = P_o \exp\{A_{o,1}t\} = \begin{bmatrix} 1 & t/2 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 0 \end{bmatrix}$$

Notice the term $t/2$ in the behavior of the system at t/ϵ . It is this unstable entry which when seen at time scale t/ϵ^2 will diverge.

To see that the limit

$$\lim_{\epsilon \downarrow 0} \exp\{A_o(\epsilon)t/\epsilon^2\} \quad (4.4.62)$$

does not exist, let us look at the entry (1,2) of $\exp\{A_0(\varepsilon)t/\varepsilon^2\}$:

$$\left(\exp\{A_0(\varepsilon)t/\varepsilon^2\}\right)_{1,2} = \left[\frac{\varepsilon}{\lambda_2} \left(e^{\lambda_2 t/\varepsilon^2} - 1 \right) - \frac{\varepsilon}{\lambda_1} \left(e^{\lambda_1 t/\varepsilon^2} - 1 \right) \right] \frac{1}{\lambda_2 - \lambda_1} \quad (4.4.63)$$

Because $\lambda_2 = -\varepsilon^2 + o(\varepsilon^2)$ the first term in (4.4.63) is of order $1/\varepsilon$ as $\varepsilon \rightarrow 0$ and therefore (4.4.62) diverges. Thus, (4.4.61) does not have well defined behavior at time scale t/ε^2 even though it has a real negative eigenvalue of order ε^2 . This is so because

$$\|\exp\{A_0(\varepsilon)t/\varepsilon^2\}\| \rightarrow \infty \quad \text{as} \quad \varepsilon \rightarrow 0$$

This behavior does not contradict the stability properties of $A_0(\varepsilon)$ because even though for every $\varepsilon \in [0,1]$

$$\sup_{t \geq 0} \|\exp\{A_0(\varepsilon)t\}\| = K(\varepsilon) < \infty \quad \forall t \geq 0$$

the upper bound $K(\varepsilon) \rightarrow \infty$ as $\varepsilon \rightarrow 0$. This example illustrates one reason why even systems which are semistable for $\varepsilon \in [0, \varepsilon_0]$ may fail to have well defined behavior at some time scales. The next example illustrates another such reason. □

Example 4.4.8

Consider the matrix

$$A_o(\epsilon) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -\epsilon^2 & \epsilon \\ 0 & -\epsilon & -\epsilon^2 \end{bmatrix}$$

It is semistable for $\epsilon > 0$ and it has three eigenvalues

$$\lambda_o = 2, \quad \lambda_{1,2} = -\epsilon^2 \pm i\epsilon$$

The matrices $A_{o,o}$ and $A_{1,o}$ for this example are:

$$A_{o,o} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad A_{1,o} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

and the MSST assumption is violated because $A_{1,o}$ has purely imaginary eigenvalues.

A simple calculation gives:

$$\exp\{A_o(\epsilon)t\} = \begin{bmatrix} e^{-2t} & 0 & 0 \\ 0 & e^{-\epsilon^2 t} \cos \epsilon t & e^{-\epsilon^2 t} \sin \epsilon t \\ 0 & e^{-\epsilon^2 t} \sin \epsilon t & e^{-\epsilon^2 t} \cos \epsilon t \end{bmatrix}$$

Clearly the system has well defined behavior at time scales t and t/ϵ but $\exp\{A_0(\epsilon)t/\epsilon^2\}$ does not have a limit as $\epsilon \downarrow 0$ because of the presence of terms of the type $e^{-t} \sin(t/\epsilon)$. \square

This example illustrates that the existence of slowly attenuated oscillations (reflected as purely imaginary poles in one of the matrices $A_{k,0}$) impedes the existence of well defined behavior at time scales at which these oscillations are at frequencies that increase without bound as $\epsilon \downarrow 0$ without commensurate increase in damping.

Examples 4.4.7 and 4.4.8 illustrate the two basic phenomena that occur whenever the multiple semistability condition is violated. If it is violated because zero is not a semisimple eigenvalue of, say, $A_{k,0}$, this is indicative of terms of the type $\epsilon^k t e^{-\epsilon^{k+1}t}$ which increase without bound as $\epsilon \downarrow 0$ for times of order t/ϵ^{k+1} . If, on the other hand, the semistability condition is violated because $A_{k,0}$ has semisimple purely imaginary eigenvalues, then the system does not have well defined behavior at time scale t/ϵ^{k+1} because of terms of the type $e^{-\epsilon^{k+1}t} \sin(\epsilon^k t)$. The only other possibility, that $A_{k,0}$ has multiple, non-semisimple, purely-imaginary eigenvalues exhibits features of the two basic cases described above. With terms of the type $\epsilon^k t e^{-\epsilon^{k+1}t} \sin(\epsilon^k t)$, the system has growing oscillations of increasing frequency at time scale t/ϵ^{k+1} .

The next theorem establishes the fact that for an arbitrary $A_0(\epsilon)$, MSST is a necessary condition for $\exp\{A_0(\epsilon)t\}$ to have well defined behavior at all time scales.

Theorem 4.4.9

Let $A_0(\epsilon)$, $\epsilon \in [0, \epsilon_0]$, be a semistable matrix with a series expansion in powers of ϵ and let $A_{k,0}$, $k \geq 0$, be the sequence of matrices constructed in Section 4.4.1.

If $A_{0,0}, A_{1,0}, \dots, A_{\ell-1,0}$ are semistable but $A_{\ell,0}$ is not, then the limit

$$\lim_{\epsilon \downarrow 0} \sup_{t > 0} \exp\{A_0(\epsilon)t/\epsilon^q\}$$

does not exist for any $\ell < q < \ell+1$. Furthermore, if $A_{\ell,0}$ has a pole on the imaginary axis (zero included) which is not semisimple then:

$$\lim_{\epsilon \downarrow 0} \sup_{t > 0} \|\exp\{A_0(\epsilon)t\}\| = \infty$$

Remarks

(1) Notice that even though in general a perturbed matrix $A_0(\epsilon)$ may have eigenvalues with expansions in fractional powers of ϵ , no singularly perturbed linear system can have a fundamental time scale at time scale t/ϵ^q for q non-integer. This is so because eigenvalues in fractional powers of ϵ may occur only if at some stage, $A_{k,0}$ does not have SSNS. Thus, eigenvalues in fractional powers of ϵ go hand in hand with unbounded amplitudes (as $\epsilon \downarrow 0$).

(2) The proof of Theorem 4.4.9 is long and it brakes the flow of ideas exposed. The reader may want to skip it on a first reading. If so, please move to page .

Proof:

We first construct the proof for $l=0$ by contradiction. Suppose that the limit:

$$\lim_{\varepsilon \downarrow 0} \exp\{A_0(\varepsilon)t/\varepsilon^q\} \quad (4.4.64)$$

exists $t > 0$ and some $0 < q < 1$. If this limit exists, so does the limit of $\exp\{P_0(\varepsilon)A_0(\varepsilon)t/\varepsilon^q\}$ as $\varepsilon \downarrow 0$ because:

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} P_0(\varepsilon) \exp\{A_0(\varepsilon)t/\varepsilon^q\} &= \\ & \lim_{\varepsilon \downarrow 0} \exp\{P_0(\varepsilon)A_0(\varepsilon)t/\varepsilon^q\} - Q_0 \end{aligned}$$

Define

$$F_0(\varepsilon) = \frac{P_0(\varepsilon)A_0(\varepsilon)}{\varepsilon^q}$$

The next step is to prove that $\sigma(F_0(\varepsilon))$ remains bounded as $\varepsilon \downarrow 0$.

Take any $0 \neq \lambda(\varepsilon) \in \sigma(P_0(\varepsilon)A_0(\varepsilon))$ and let $\phi(\varepsilon)$ be an eigenvector with $\lambda(\varepsilon)$ normalized so that $\|\phi(\varepsilon)\| = 1$. Then,

$$\begin{aligned} \exp\{P_0(\varepsilon)A_0(\varepsilon)t/\varepsilon^q\}\phi(\varepsilon) &= \\ & \exp\{\operatorname{Re} \lambda(\varepsilon)t/\varepsilon^q\} \cdot \exp\{i \operatorname{Im} \lambda(\varepsilon)t/\varepsilon^q\} \cdot \phi(\varepsilon) \end{aligned}$$

and, if $\varepsilon_m \rightarrow 0$ is a sequence for which $\phi(\varepsilon_m)$ converges, then

$$\exp\{\operatorname{Re} \lambda(\varepsilon_m)t/\varepsilon_m^q\}, \exp\{i \operatorname{Im} \lambda(\varepsilon_m)t/\varepsilon_m^q\} \cdot \phi(\varepsilon_m) \quad (4.4.65)$$

must also converge as $m \rightarrow \infty$. Notice that because the trace $A_0(\varepsilon)$ has a series expansion in integer powers of ε and all its eigenvalues, $\mu(\varepsilon)$, have non-positive real parts there can be no cancellations and therefore,

$$\operatorname{Re} \mu(\varepsilon) = O_s(\varepsilon^k)$$

for some integer $k \geq 0$. Thus, from $\lambda(\varepsilon) \rightarrow 0$, we conclude

$$\frac{\operatorname{Re} \lambda(\varepsilon_m)}{\varepsilon_m^q} \rightarrow 0 \quad \text{as } m \rightarrow \infty$$

From the fact that (4.4.65) converges it follows that $\operatorname{Im} \lambda(\varepsilon_m)/\varepsilon_m^q$ must also converge as $m \rightarrow \infty$. Thus $\sigma(F_0(\varepsilon))$ remains bounded as $\varepsilon \rightarrow 0$ and it is therefore possible to choose t_1 such that

$$|\operatorname{Im} \sigma(F_0(\varepsilon)t_1)| < \pi$$

for ε small enough. Now, if Ln denotes the principal branch of the logarithmic function, we obtain:

$$\operatorname{Ln} \exp\{F_0(\varepsilon)t_1\} = F_0(\varepsilon)t_1 = \left(\frac{D_0}{\varepsilon^q} + G_0(\varepsilon) \right) t_1 \quad (4.4.66)$$

The last equality following from Lemma 3.4.2 with D_0 being the eigennilpotent for the zero eigenvalue of $A_{0,0}$ and $G_0(\varepsilon)$ a function of ε , continuous at $\varepsilon=0$. The limit

$$\lim_{\varepsilon \rightarrow 0} \operatorname{Ln} \exp\{F_0(\varepsilon)t_1\} \stackrel{\Delta}{=} B(t_1)$$

is well defined because of the boundedness of $\sigma(F_0(\epsilon)t_1)$. Now, taking limits as $\epsilon \rightarrow 0$ in (4.4.66) we get a contradiction unless $D_0 = 0$. We thus conclude that for the limit (4.4.64) to exist, $A_{0,0}$ must have SSNS.

Suppose now that $A_{0,0}$ has some purely imaginary eigenvalue μ . Then there exists an eigenvalue $\mu(\epsilon)$ such that $\mu(\epsilon) \rightarrow \mu$ as $\epsilon \rightarrow 0$. Let $\phi(\epsilon)$ be an eigenvector with eigenvalue $\mu(\epsilon)$ and $\|\phi(\epsilon)\| = 1$. Pick a sequence $\epsilon_m \rightarrow 0$ for which $\phi(\epsilon_m)$ converges as $m \rightarrow \infty$. Then, if (4.4.64) converges so must do

$$\phi(\epsilon_m)^T \exp\{A_{0,0}(\epsilon_m)t/\epsilon_m^q\} \phi(\epsilon_m) = \exp\{\mu(\epsilon_m)t/\epsilon_m^q\}$$

which is a contradiction. We have thus shown that if $A_{0,0}$ is not semistable then (4.4.64) does not exist.

To prove the theorem for an arbitrary l notice that by the semistability property of $A_{0,0}, \dots, A_{l-1,0}$, we can construct the matrices $A_k(\epsilon)$, $k=0, \dots, l$, as in Section 4.4.1, and using the same algebraic manipulations as in the proof of Theorem 4.4.4 (see eq. (4.4.34)), we can write:

$$\begin{aligned} \exp\{A_0(\epsilon)t/\epsilon^q\} &= \exp\{A_l(\epsilon)t/\epsilon^{q-l}\} - \sum_{k=0}^{l-1} Q_k(\epsilon) + \\ &+ \sum_{k=0}^{l-1} Q_k(\epsilon) \exp\{A_k(\epsilon)t/\epsilon^{q-k}\} \end{aligned} \quad (4.4.65)$$

Suppose now that

$$\lim_{\epsilon \rightarrow 0} \exp\{A_0(\epsilon)t/\epsilon^q\} \quad (4.4.66)$$

exists for all $t > 0$ and some $l < q \leq l+1$. Using (4.4.65) it is easy to see that this implies that

$$\lim_{\varepsilon \downarrow 0} \exp\{A_\ell(\varepsilon)t/\varepsilon^{q-l}\} \quad l < q \leq l+1 \quad (4.4.67)$$

also exists (the second and third terms in (4.4.65) have well defined limits because $A_{k,0}$, $k=0, \dots, l-1$ are semistable). But (4.4.67) is the same kind of limit as (4.4.64) and we have already seen that its existence implies $A_{l,0}$ semistable which is again a contradiction indicating that, under the conditions of the theorem, (4.4.66) cannot converge as $\varepsilon \downarrow 0$.

To prove the second part of the theorem suppose that $A_{l,0}$ has an eigenvalue on the imaginary axis which is not semisimple. Then $\forall M < \infty$ there exists a $T < \infty$ such that

$$\|\exp\{A_{l,0}T\}\| > M$$

and because

$$\|\exp\{A_\ell(\varepsilon)T\} - \exp\{A_{l,0}T\}\| \rightarrow 0 \quad \text{as } \varepsilon \downarrow 0$$

we conclude that

$$\|\exp\{A_\ell(\varepsilon)T\}\| > M/2$$

for ε small enough.

Thus,

$$\lim_{\varepsilon \downarrow 0} \sup_{t \geq 0} \|\exp\{A_\ell(\varepsilon)t\}\| = \infty \quad (4.4.68)$$

From eq. (4.4.65) we have:

$$\begin{aligned} \sup_{t \geq 0} \|\exp\{A_\ell(\varepsilon)t\}\| &= \sup_{t \geq 0} \|\exp\{A_\ell(\varepsilon)\varepsilon^\ell t\}\| \leq \\ &\sup_{t \geq 0} \|\exp\{A_0(\varepsilon)t\}\| + \sum_{k=0}^{\ell-1} \|Q_k(\varepsilon)\| + \\ &\sup_{t \geq 0} \sum_{k=0}^{\ell-1} \|Q_k(\varepsilon)\exp\{A_k(\varepsilon)\varepsilon^k t\}\| \end{aligned} \quad (4.4.69)$$

and by semistability of $A_{k,0}$ $k=0, \dots, \ell-1$, it follows from (4.4.41)

that:

$$\sup_{t \geq 0} \|Q_k(\varepsilon)\exp\{A_k(\varepsilon)\varepsilon^k t\}\| \leq M_k < \infty \quad (4.4.70)$$

Taking limits as $\varepsilon \downarrow 0$ in (4.4.69) and using (4.4.70) and (4.4.68)

we get the desired result:

$$\lim_{\varepsilon \downarrow 0} \sup_{t \geq 0} \|\exp\{A_0(\varepsilon)t\}\| = \infty \quad \square$$

The results in Sections 4.4.2 and 4.4.3 together with Theorem 4.4.9 indicate that multiple semistability is a *necessary and sufficient*

condition for a system to have a well defined multiple time scale behavior. From a modeling viewpoint it implies that if the system matrix of a linear time-invariant system has small entries that result in nearly decoupled evolutions taking place at different time scales then, proper modelling will result in a singular perturbed system satisfying the MSST condition. Also, as we will see in Chapter V, there are certain classes of systems which *always* satisfy this condition. In general, however, the dependence of the system's matrix on a small parameter will more likely be based on physical considerations about several parameters of the system. In this case, one needs to compute the matrices $A_{k,0}$ to determine whether a time scale decomposition of the system is feasible. In the next section we address this aspect of the problem.

4.4.5 Computation of the reduced-order models

It follows from our development in previous sections that the matrices $A_{k,0}$ play a fundamental role in the asymptotic analysis of singularly perturbed systems. We now focus on an algorithm for the computation of these matrices.

It is convenient to think of the sequence of matrices $A_k(\epsilon)$, $k=1, \dots, m$, constructed in Section 4.4.1 as defining a rectangular array of matrices A_{ij} , $i=0, 1, \dots, m$, $j \geq 0$, as shown in Fig. 4.2. By definition of the matrices $A_k(\epsilon)$, it follows from Lemma 3.4.2 and eq. (2.4.29)

$$\begin{array}{ccccccc} A_0(\epsilon) & & A_{00} & A_{01} & A_{02} & \cdots & A_{0m} & \cdots \\ A_1(\epsilon) & & A_{10} & A_{11} & \cdots & \cdots & A_{1,m-1} & \cdots \\ A_2(\epsilon) & & A_{20} & \cdots & \cdots & \cdots & A_{2,m-2} & \cdots \\ \vdots & & \vdots & & & & & \\ \vdots & & \vdots & & & & & \\ A_m(\epsilon) & & A_{m0} & \cdots & \cdots & \cdots & & \end{array}$$

Figure 4.2: The array of matrices A_{ij} .

that the $(i+1)^{\text{th}}$ row in Fig. 4.2 can be computed from the i^{th} row using the formula:

$$A_{i+k,j} = - \sum_{p=1}^{j+1} (-1)^p \sum_{\substack{v_1+\dots+v_p=j+1 \\ k_1+\dots+k_{p+1}=p-1 \\ v_i \geq 1, k_j \geq 0}} S_i^{(k_1)} A_{i,v_1} S_i^{(k_2)} \dots A_{i,v_p} S_i^{(k_{p+1})} \quad (4.4.71)$$

$$\begin{aligned} i &= 0, 1, \dots, m \\ j &\geq 0 \end{aligned}$$

where

$$S_i^{(0)} = -P_i$$

$$S_i^{(k)} = (A_{i,o}^\#)^k \quad k > 0$$

(remember, P_i is the projection on $N(A_{i,o})$ along $R(A_{i,o})$ and $A_{i,o}^\#$ denotes the generalized inverse of $A_{i,o}$). Notice that the structure of (4.4.71) permits us to grow the array $A_{i,j}$ triangularly: A_{10} is computed from A_{00} and A_{01} ; A_{20} requires A_{10} and A_{11} which in turn involve A_{00} , A_{01} and A_{02} ; in general, to compute the first column up to A_{k0} involves the matrices A_{ij} , $i=0, \dots, k$, $j=0, \dots, k-i$. As we have already seen in Proposition 4.4.2, only a finite number of matrices A_{ko} need to be computed and it thus follows from (4.4.71) that to do so requires only a finite amount of computational effort.

Although algorithm (4.4.71) is attractive for its recursive nature, a closer look at the structure of (4.4.71) reveals that it involves a large number of superflows computations. Without addressing the general issue of the most efficient way to compute the matrices A_{ko} , we will now give an explicit expression for the matrices A_{00} , A_{10} , A_{20} and A_{30} in terms of the first row in Fig. 4.2 using formula (4.4.71) and will conjecture a simplified version of (4.4.71) that excludes terms which are cancelled in the course of the recursion.

To proceed with the computation of A_{ko} for $k=0,1,\dots,3$, use (4.4.71) to write:

$$A_{i+1,0} = P_i A_{i1} P_i \quad (4.4.72)$$

$$A_{i+1,1} = P_i A_{i2} P_i - P_i A_{i1} P_i A_{i0}^\# - A_{i0}^\# A_{i1} P_i A_{i1} P_i - P_i A_{i0}^\# A_{i0}^\# A_{i1} P_i \quad (4.4.73)$$

$$A_{i+1,2} = P_i A_{i3} P_i - P_i A_{i1} P_i A_{i2} A_{i0}^\# - A_{i0}^\# A_{i1} P_i A_{i2} P_i - P_i A_{i0}^\# A_{i0}^\# A_{i2} P_i - P_i A_{i2} P_i A_{i1} A_{i0}^\# - A_{i0}^\# A_{i2} P_i A_{i1} P_i - P_i A_{i2} A_{i0}^\# A_{i0}^\# A_{i1} P_i + P_i A_{i1} P_i A_{i1} A_{i0}^\# + P_i A_{i0}^\# A_{i0}^\# A_{i1} P_i A_{i1} A_{i0}^\# + A_{i0}^\# A_{i1} P_i A_{i1} P_i A_{i1} A_{i0}^\# + A_{i0}^\# A_{i1} P_i A_{i0}^\# A_{i0}^\# A_{i1} P_i + A_{i0}^\# A_{i0}^\# A_{i1} P_i A_{i1} P_i + P_i A_{i0}^\# A_{i0}^\# A_{i0}^\# A_{i1} P_i - P_i A_{i1} P_i A_{i1} P_i A_{i0}^\# (A_{i0}^\#)^2 - P_i A_{i1} P_i A_{i1} (A_{i0}^\#)^2 A_{i0}^\# A_{i1} P_i - P_i A_{i1} (A_{i0}^\#)^2 A_{i0}^\# A_{i1} P_i A_{i1} P_i - (A_{i0}^\#)^2 A_{i0}^\# A_{i1} P_i A_{i1} P_i A_{i1} P_i \quad (4.4.74)$$

To simplify these expressions remember that the direct sum decomposition

$$R^n = R(A_{o0}) \oplus \dots \oplus R(A_{m0}) \oplus \left(\sum_{k=0}^m N(A_{ko}) \right)$$

together with the SSNS of A_{ko} imply

$$N(A_{ko}) = R(A_{o0}) \oplus \dots \oplus R(A_{k-1,o}) \oplus R(A_{k+1,o}) \oplus \dots \oplus R(A_{m0}) \oplus \left(\sum_{k=0}^m N(A_{ko}) \right)$$

and therefore,

$$P_i P_j = P_j P_i \tag{4.4.75}$$

$$P_i A_{jo} = A_{jo} P_i = P_i A_{jo} P_i = (1 - \delta_{ij}) A_{jo} \tag{4.4.76}$$

On the other hand, by Lemma 3.2.7,

$$P_i A_{io}^\# = A_{io}^\# P_i = 0 \tag{4.4.77}$$

We now use these properties to simplify (4.4.73) and (4.4.74). Using (4.4.76) and (4.4.78) in (4.4.73) and (4.4.74) we get:

$$A_{i+1,1} = P_i (A_{i2} - A_{il} A_{io}^\# A_{il}) P_i \tag{4.4.78}$$

$$A_{i+1,2} = P_i (A_{i3} - A_{il} A_{io}^\# A_{i2} - A_{i2} A_{io}^\# A_{il} + A_{il} A_{io}^\# A_{il} A_{io}^\# A_{il} - A_{il} P_i A_{il} (A_{io}^\#)^2 A_{il} - A_{il} (A_{io}^\#)^2 A_{il} P_i A_{il}) P_i \tag{4.4.79}$$

Still more cancellations occur when these expressions are used recursively to compute the matrices A_{ko} as we will now see.

Proposition 4.4.10

Let $A_o(\epsilon)$ be given by

$$A_o(\epsilon) = A_{oo} + \epsilon A_{o1} + \epsilon^2 A_{o2} + \epsilon^3 A_{o3} + o(\epsilon^3)$$

then the matrices A_{ko} , $k=0,1,2,3$, are given by:

$$A_{oo}$$

$$A_{10} = P_o A_{o1} P_o$$

$$A_{20} = P_1 P_o (A_{o2} - A_{o1} A_{oo}^\# A_{o1}) P_o P_1$$

$$A_{30} = P_2 P_1 P_o (A_{o3} - A_{o1} A_{oo}^\# A_{o2} - A_{o2} A_{oo}^\# A_{o1} +$$

$$A_{o1} A_{oo}^\# A_{o1} A_{oo}^\# A_{o1} - A_{o2} A_{10}^\# A_{o2} +$$

$$A_{o2} A_{10}^\# A_{o1} A_{oo}^\# A_{o1} - A_{o1} A_{oo}^\# A_{o1} A_{10}^\# A_{o2} -$$

$$A_{o1} A_{oo}^\# A_{o1} A_{10}^\# A_{oo}^\# A_{o1}) P_o P_1 P_2$$

(4.4.83)

Proof:

Equation (4.4.81) follows directly from (4.4.72). To prove (4.4.82) use (4.4.72) to write

$$A_{20} = P_1 A_{11} P_1$$

and now substitute A_{11} using (4.4.79) with $i=0$:

$$A_{20} = P_1 P_o (A_{o2} - A_{o1} A_{oo}^\# A_{o1}) P_o P_1$$

To prove (4.4.83) use (4.4.72) again with $i=2$ and (4.4.79)

with $i=1$:

$$\begin{aligned} A_{30} &= P_2 A_{21} P_2 \\ &= P_2 P_1 (A_{12} - A_{11} A_{10}^{\#} A_{11}) P_1 P_2 \end{aligned} \quad (4.4.84)$$

Let us now compute the first term in (4.4.84) using (4.4.80) with

$i=0$,

$$\begin{aligned} P_2 P_1 A_{12} P_1 P_2 &= P_2 P_1 P_0 (A_{03} - A_{01} A_{00}^{\#} A_{02} - A_{02} A_{00}^{\#} A_{01} + \\ &A_{01} A_{00}^{\#} A_{01} A_{00}^{\#} A_{01} - A_{01} P_0 A_{01} (A_{00}^{\#})^2 A_{01} - A_{01} (A_{00}^{\#})^2 A_{01} P_0 A_{01}) P_0 P_1 P_2 \end{aligned} \quad (4.4.85)$$

We can simplify this expression by noting that

$$P_1 A_{10} = P_1 P_0 A_{01} P_0 = P_0 A_{01} P_0 P_1 = 0$$

resulting in

$$\begin{aligned} P_2 P_1 A_{12} P_1 P_2 &= P_2 P_1 P_0 (A_{03} - A_{01} A_{00}^{\#} A_{02} - A_{02} A_{00}^{\#} A_{01} + \\ &A_{01} A_{00}^{\#} A_{01} A_{00}^{\#} A_{01}) P_0 P_1 P_2 \end{aligned} \quad (4.4.86)$$

To compute the second term in (4.4.84) substitute A_{11} by (4.4.79)

with $i=0$. This together with (4.4.86) gives (4.4.83). □

Equations (4.4.81)-(4.4.83) give an explicit expression for the first four reduced-order models in terms of the given data $A_{o,k}$.

To obtain explicit expressions for reduced-order models beyond the fourth time scale, a huge amount of algebraic manipulation is required. Of course, it is always possible to use the recursive formula (4.4.71) for specific examples, but the computational cost of this method, seems to be quite high. We now proceed to describe a conjectured, simplified version of (4.4.71).

By comparing (4.4.79)-(4.4.80) to (4.4.73)-(4.4.74) it is clear that a large number of terms in (4.4.71) can be immediately discarded because they get cancelled at one stage or another of the recursion. Furthermore, notice by comparing (4.4.80) to (4.4.81)-(4.4.83) that no term containing powers of a group inverse appears in the final result. The formula,

$$\bar{A}_{i+1,j} = - \sum_{p=1}^{j+1} (-1)^p v_1 + \dots + v_p = j+1 \bar{P}_i \bar{A}_{i,v_1} \bar{A}_{i,o}^{\#} \dots \bar{A}_{i,o}^{\#} \bar{A}_{i,v_p} \bar{P}_i \quad (4.4.87)$$

does not contain any of the terms in (4.4.73) or (4.4.74) that get cancelled in the recursion process and in fact results in the same expression (4.4.81)-(4.4.83) for $A_{ko} = \bar{A}_{ko}$, $k=1,2,3$. This together with the arguments given in [Del 82], in the context of Markov chains, using a totally different approach lead us to the following.

Conjecture 4.4.11

The recursion formulas (4.4.71) and (4.4.87) result in the same matrices in the first column of Table 4.2, i.e.,

$$A_{ko} = \bar{A}_{ko} \quad k=0,1,\dots,m$$

Remark: Notice that the conjecture refers only to the first column of Table 4.2 but certainly not to the other matrices. It is quite clear that formula (4.4.87) cannot be used to compute the matrices $A_k(\epsilon)$ appearing in the *exact* time scale decomposition of $\exp\{A_0(\epsilon)t\}$ (i.e., in Theorem 4.4.4). It seems plausible, however, that both formulas will produce the same asymptotic approximation. (For sure in systems with four or less time scales).

There are certainly more issues to be explored before the computational complexity of the multiple time scale and aggregation methods proposed in this thesis can be ascertained. This aspect of the problem remains open and beyond the scope of this thesis.

To conclude this section we derive simple upper bounds for the number of fundamental time scales, i.e., the number of non-zero matrices in $\{A_{k,0}\}_{k=0}^m$, and for the slowest time scale, that is, t/ϵ^m . Remember that under the MSST condition, the sequence $A_{k,0}$ ends at some $k=m$ for which

$$\sum_{k=0}^m \text{rank } A_{k,0} = n \text{rank } A_0(\epsilon)$$

or

$$\sum_{k=1}^m \text{rank } A_{k,0} = \text{rank } A_0(\epsilon) - \text{rank } A_0(0) \quad (4.4.88)$$

$\epsilon \in (0, \epsilon_0]$

and therefore the maximum number of fundamental time scales in addition to time scale t equals the size of the rank discontinuity of $A_0(\epsilon)$ at $\epsilon=0$. To find an upper bound for m consider first the linear perturbation case, i.e., let $A_0(\epsilon)$ be given by:

$$A_0(\epsilon) = A_{00} + \epsilon A_{01}$$

The eigenvalues of $A_0(\epsilon) = A_{00} + \epsilon A_{01}$ are the solution of a polynomial of degree n with coefficients that are themselves polynomials in ϵ of degree $\leq n$. A simple argument shows that there can be no eigenvalue, $\lambda(\epsilon) \neq 0$, of $A_0(\epsilon)$ such that $\lambda(\epsilon) = o(\epsilon^n)$. In effect, let

$$s^n + p_{n-1}(\epsilon)s^{n-1} + \dots + p_1(\epsilon)s + p_0(\epsilon) = 0$$

be the characteristic polynomial of $A_{00} + \epsilon A_{01}$. The coefficients $p_{n-i}(\epsilon)$ are polynomials in ϵ of degree $\leq i$. Then,

$$\lambda(\epsilon)^{n-1} + p_{n-1}(\epsilon)\lambda(\epsilon)^{n-2} + \dots + p_1(\epsilon) + \frac{p_0(\epsilon)}{\lambda(\epsilon)} = 0 \quad (4.4.89)$$

If $\lambda(\epsilon)$ were of order $o(\epsilon^n)$ and $p_0(\epsilon) \neq 0$ then $p_0(\epsilon)/\lambda(\epsilon) \rightarrow \infty$ as $\epsilon \rightarrow 0$ and (4.4.89) cannot be satisfied. If $p_0(\epsilon) = 0$ then the same can be set about

$$\lambda(\epsilon)^{n-2} + p_{n-1}(\epsilon)\lambda(\epsilon)^{n-3} + \dots + p_2(\epsilon) + \frac{p_1(\epsilon)}{\lambda(\epsilon)} = 0 \quad (4.4.90)$$

which cannot be satisfied unless $p_1(\epsilon) = 0$. Proceeding in this way it is concluded that if $\lambda(\epsilon) = o(\epsilon^n)$ then $\lambda(\epsilon) = 0$. It then follows

from Proposition 4.4.3 -iii) that $A_k(\epsilon) = 0$ for $k > n$ and therefore $m \leq n$. Similarly, in the case of a non-linear perturbation of finite order,

$$A_o(\epsilon) = \sum_{j=0}^p A_{oj} \epsilon^j$$

we have $m \leq n \cdot p$.

4.5 Partial Time Scale Decomposition

In this section we analyze the multiple time scale behavior of semistable singularly perturbed systems that *do not* satisfy the MSST condition. In general, these systems have well defined behavior at some time scales but not at all time scales, and their behavior over the infinite time interval $[0, \infty)$ cannot be reconstructed from their evolutions at different time scales. It may still be useful, however, to isolate the time scales at which they have well defined behavior and to compute their evolutions at these time scales. This is the problem we address here.

Let $A_o(\epsilon)$ be the system matrix under consideration and let A_{ko} , $k=0,1,\dots,m$ be the sequence of matrices derived as in Section 4.4.1. To analyze the different possibilities that may occur if the MSST condition is violated, we distinguish the following cases:

Case 1:

The matrices A_{k0} , $k=0,1,\dots,m$, satisfy the MSSNS condition. All of them, except $A_{\ell 0}$, are semistable but $A_{\ell 0}$ violates the semistability condition because it has some purely imaginary eigenvalues. The purely imaginary eigenvalues are semisimple. For reasons that will become clear later, we will refer to this case as the uniform stability case.

Case 2:

As Case 1 but some purely imaginary eigenvalue of $A_{\ell 0}$ is not semisimple.

Case 3:

The matrix A_{m0} does not have SSNS.

These three cases indicate the basic possibilities that may occur when the MSST condition is violated. In a given matrix $A_0(\epsilon)$, any combination of them may occur at different stages of the sequence A_{k0} , but the analysis of this general case can be decomposed into the single feature cases described above.

The results in Section 4.4 can be extended with minor modifications only to the uniform stability case and this is done in the next section. In Section 4.5.2 we briefly discuss the difficulties encountered in Cases 2 and 3.

4.5.1 Uniformly Stable Systems

To provide some motivation for the result of this section, consider a matrix $A_0(\epsilon)$ which is semistable for $\epsilon \in [0, \epsilon_0]$. Semistability implies uniform boundedness of $\exp\{A_0(\epsilon)t\}$ with respect to t , i.e.,

$$\sup_{t \geq 0} \|\exp\{A_0(\epsilon)t\}\| = K(\epsilon) < \infty \quad (4.5.1)$$

But, as shown by Example 4.4.7, in singularly perturbed systems $K(\epsilon)$ may become unbounded as $\epsilon \downarrow 0$. This kind of behavior indicates that in some systems the ϵ -dependence, in addition to generating eigenvalues of different orders of magnitude in ϵ (that is, different time scales), also produces amplitudes that become unbounded as $\epsilon \downarrow 0$.

As we have seen in Section 4.4.4, the presence of increasingly large amplitudes as $\epsilon \downarrow 0$ impedes the complete multiple time scale analysis of these systems. Here we analyze the multiple time scale behavior of systems in which the ϵ -dependence does not give rise to unbounded amplitudes as $\epsilon \downarrow 0$.

We will say that a stable system with system matrix $A_0(\epsilon)$ is *uniformly stable in ϵ* or, in short, that $A_0(\epsilon)$ satisfies the *uniform stability (US) condition* if:

$$\|\exp\{A_0(\epsilon)t\}\| \leq K \quad \forall t \geq 0 \quad \forall \epsilon \in [0, \epsilon_0]$$

for some $K > 0$ independent of ϵ . The following Proposition establishes a link between the US condition and the spectrum of the matrices A_{k0} $k=0,1,\dots,m$ derived from $A_0(\epsilon)$.

Proposition 4.5.1

If $A_0(\epsilon)$ is uniformly stable then it also satisfies the MSSNS condition and if any of the reduced-order models A_{k0} has eigenvalues on the imaginary axis they must be semisimple.

Proof:

From Prop. 4.2.1. we have

$$\exp\{A_0(\epsilon)t\} \rightarrow \exp\{A_0(0)t\} \text{ as } \epsilon \rightarrow 0, \forall t \geq 0$$

and therefore

$$\|\exp\{A_0(\epsilon)t\}\| \leq K$$

implies

$$\|\exp\{A_{00}t\}\| \leq K$$

which indicates that all eigenvalues of A_{00} on the imaginary axis (zero included) must be semisimple. Doing the same manipulation as in (4.4.33) we get:

$$\exp\{A_1(\epsilon)t\} = \exp\{A_0(\epsilon)t/\epsilon\} + Q_0(\epsilon) - Q_0(\epsilon)\exp\{A_0(\epsilon)t/\epsilon\}$$

and uniform stability of $A_0(\epsilon)$ plus continuity of $Q_0(\epsilon)$ imply that $A_1(\epsilon)$ is U.S., i.e.

$$\exp\{A_1(\epsilon)t\} \leq K_1$$

As before this implies that all imaginary axis poles of A_{10} are semi-simple and the same process can be repeated to prove that all A_{k0} 's satisfy this property. □

Uniform stability guarantees MSSNS but not MSST because some of the matrices A_{k0} may have purely imaginary eigenvalues. However, MSST implies US.

Uniformly stable systems may fail to have well defined behavior at certain time scales because of the presence of oscillatory modes in some of the reduced-order models A_{k0} . These oscillations become of infinite frequency when seen at slower time scales. It is important to notice that the appearance of such unattenuated oscillations in some of the reduced-order models does not necessarily imply that the matrix $A_0(\epsilon)$ has some purely imaginary eigenvalues. Instead, they could as well correspond to eigenvalues with a negative real part that converges to zero faster than its imaginary part. As, for example, in modes like $e^{-\epsilon^2 t} \sin \epsilon t$. This mode is seen as purely oscillatory at time scale t/ϵ and when the attenuation effects are beginning to be felt, at time scale t/ϵ^2 , the oscillations become of infinite frequency.

To avoid this lack of well defined limit the oscillatory modes must be excluded before analyzing the system at slower time scales. The following Proposition indicates how can this be accomplish.

Proposition 4.5.2

Let $A_0(\epsilon)$ be uniformly stable and suppose that all the reduced order models A_{k0} , $k=0,1,\dots,m$, are semi-stable except A_{l0} . Then:

i)

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \|P_l \exp\{A_0(\epsilon)t\} - \phi^l(\epsilon, t)\| = 0 \quad (4.5.2)$$

where

$$\phi^l(\epsilon, t) = \sum_{\substack{k=0 \\ k \neq l}}^m Q_k \exp\{A_{k0} \epsilon^k t\} + P_0 P_1 \dots P_m \quad (4.5.3)$$

ii)

$$\lim_{\epsilon \downarrow 0} \sup_{\delta \leq t \leq T} \|\exp\{A_0(\epsilon)t/\epsilon^k\} - \Phi_k(t)\| = 0 \quad (4.5.4)$$

$$\forall \delta > 0, \forall T < \infty \\ k=0,1,\dots,l$$

$$\lim_{\epsilon \downarrow 0} \sup_{\delta \leq t \leq T} \|P_l \exp\{A_0(\epsilon)t/\epsilon^k\} - \Phi_k(t)\| = 0 \quad (4.5.5)$$

$$\forall \delta > 0, \forall T < \infty \\ k=l+1,\dots,m$$

where:

$$\begin{aligned} \Phi_k(t) &= Q_k \exp\{A_{k0} t\} + P_0 \dots P_k \\ &= P_0 \dots P_{k-1} \exp\{A_{k0} t\} \end{aligned}$$

Proof:

$$\|P_l \exp\{A_0(\epsilon)t\} - \phi^l(\epsilon, t)\| \leq \quad (4.5.6)$$

$$\|P_l(\epsilon) \exp\{A_0(\epsilon)t\} - \phi^l(\epsilon, t)\| + \|(P_l(\epsilon) - P_l) \exp\{A_0(\epsilon)t\}\|$$

From the proof of Corollary 4.5.1 it follows that

$$\lim_{\varepsilon \downarrow 0} \sup_{t \geq 0} \| P_\ell(\varepsilon) \exp\{A_\circ(\varepsilon)t\} - \phi^\ell(\varepsilon, t) \| = 0$$

and because $A_\circ(\varepsilon)$ is US,

$$\lim_{\varepsilon \downarrow 0} \sup_{t \geq 0} \| (P_\ell(\varepsilon) - P_\ell) \exp\{A_\circ(\varepsilon)t\} \| \leq$$

$$\lim_{\varepsilon \downarrow 0} K \| P_\ell(\varepsilon) - P_\ell \| = 0$$

proving (4.5.2). From (4.5.2), (4.5.3) and (4.4.50) it follows that

$$\begin{aligned} P_\ell \exp\{A_\circ(\varepsilon)t/\varepsilon^k\} &= Q_k \exp\{A_{k0}t\} + \sum_{p=k+1}^m Q_p + P_\circ \dots P_m + o(1) \\ &= Q_k \exp\{A_{k0}t\} + P_\circ \dots P_k + o(1) \end{aligned}$$

uniformly on $[\delta, T]$ proving (4.5.5). The proof of (4.5.4) has already been given in Proposition 4.5.1. □

It is a simple matter to extend the above proposition to the case in which several of the matrices A_{k0} fail to be semistable. Eq. (4.5.2) still holds if $\exp\{A_\circ(\varepsilon)t\}$ is multiplied by the corresponding projections and we will now have a sequence of results similar to (4.5.4) and (4.5.5) with an increasing number of projections used as slower and slower time scales are considered.

In any of the cases indicated at the beginning of Section 4.5, one can always think of eliminating the modes that violate the MSST condition by using the correct eigenprojections to annihilate them. In general, however these projections will be given by an infinite series in powers of ϵ and it is not clear how many terms in these series are needed to keep the contribution of the undesirable modes within $o(1)$.

The fact that in the case analyzed here the projections required to annihilate the undesired modes are ϵ -independent makes Proposition 4.5.2 the most natural extension of the results derived for MSST systems. In Chapter VI we present an application of this idea to the determination of coherence areas in electric power systems.

4.5.2 Non Uniformly Stable Systems

Consider now Cases 2 and 3 specified at the beginning of Section 4.5. In both cases the matrices $A_{k,0}$ $k=0,1,\dots,\ell-1$ are assumed to be semi-stable but the MSST condition is broken by $A_{\ell,0}$ in a different way for each case. As established by Theorem 4.4.9 we have under this circumstances that:

$$\lim_{\epsilon \downarrow 0} \exp\{A_{\circ}(\epsilon)t/\epsilon^q\} \quad \ell < q \leq \ell+1 \quad (4.5.7)$$

does not exist and, furthermore, that:

$$\lim_{\epsilon \downarrow 0} \sup_{t > 0} \|\exp\{A_{\circ}(\epsilon)t\}\| = \infty \quad (4.5.8)$$

A full time scale decomposition and uniform asymptotic approximation is possible only up to time scale t/ϵ^{ℓ} , and it is given in the following proposition.

Proposition 4.5.3

Let $A_0(\epsilon)$ be such that the matrices $A_{k,0}$, $k=0,1,\dots,\ell-1$ are semistable, then:

$$\lim_{\epsilon \downarrow 0} \sup_{t \in [0, T/\epsilon^\ell]} \|\exp\{A_0(\epsilon)t\} - \phi^\ell(\epsilon, t)\| = 0 \quad (4.5.9)$$

where

$$\phi^\ell(\epsilon, t) = \sum_{k=0}^{\ell} Q_k \exp\{A_{k,0} \epsilon^k t\} + P_0 P_1 \dots P_\ell \quad (4.5.10)$$

Proof:

This is essentially a reduced version of Theorem 4.4.5 and it is proven the same way. □

Beyond time scale t/ϵ^ℓ little can be said in general terms. In Case 2, the matrix $A_0(\epsilon)$ still satisfies the MSSNS condition and the following proposition shows how, at least in principle, one can go time t/ϵ^ℓ by using the appropriate projection to annihilate the modes that produce unbounded amplitudes as $\epsilon \downarrow 0$.

Proposition 4.5.4

Let $A_0(\epsilon)$ satisfy the MSSNS condition and let $A_{k,0}$, $k=0,1,\dots,m$ be the matrices defined in Section 4.4.1.

Suppose that $A_{k,0}$, $k \neq \ell$, are semistable. Then,

i)

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \|\mathcal{P}_\ell(\epsilon) \exp\{A_0(\epsilon)t\} - \phi^\ell(\epsilon, t)\| = 0 \quad (4.5.11)$$

where

$$\phi^\ell(\varepsilon, t) = \sum_{\substack{k=0 \\ k \neq \ell}}^m Q_k \exp\{A_k \varepsilon^k t\} + P_0 P_1 \dots P_m \quad (4.5.12)$$

ii)

$$\lim_{\varepsilon \downarrow 0} \sup_{\delta < t \leq T} \|\exp\{A_0(\varepsilon)t/\varepsilon^k\} - \Phi_k(t)\| = 0 \quad (4.5.13)$$

$$\forall \delta > 0, \forall T < \infty$$

$$k=0, \dots, \ell$$

$$\lim_{\varepsilon \downarrow 0} \sup_{\delta < t \leq T} \|P_\ell(\varepsilon) \exp\{A_0(\varepsilon)t/\varepsilon^k\} - \Phi_k(t)\| = 0 \quad (4.5.14)$$

$$\forall \delta > 0, \forall T < \infty$$

$$k=\ell+1, \dots, m$$

where

$$\begin{aligned} \Phi_k(t) &= Q_k \exp\{A_{k0} t\} + P_0 \dots P_k \\ &= P_0 \dots P_{k-1} \exp\{A_{k0} t\} \end{aligned} \quad (4.5.15)$$

Proof:

From Theorem 4.4.4 and the fact that $P_\ell(\varepsilon)Q_k(\varepsilon) = Q_k(\varepsilon)$ if $k \neq \ell$ and $P_\ell(\varepsilon)Q_\ell(\varepsilon) = 0$ (see Proposition 4.4.1-iii) we have:

$$P_\ell(\varepsilon) \exp\{A_0(\varepsilon)t\} = \sum_{\substack{k=0 \\ k \neq \ell}}^m Q_k(\varepsilon) \exp\{A_k(\varepsilon)\varepsilon^k t\} + P_0(\varepsilon) \dots P_m(\varepsilon)$$

and because A_{k0} , $k \neq \ell$, are semistable it follows as in the proof of

Theorem 4.4.5 that

$$\|Q_k(\varepsilon) \exp\{A_k(\varepsilon)\varepsilon^k t\} - Q_k \exp\{A_{k0} \varepsilon^k t\}\| \leq K(\varepsilon) \rightarrow 0 \text{ as } \varepsilon \downarrow 0 \quad (4.5.16)$$

which proves (4.5.1).

By the MSSNS of $A_0(\epsilon)$ it follows from the proof of Theorem 4.4.4 (see 4.4.33 and 4.4.34) that

$$\begin{aligned} \exp\{A_0(\epsilon)t/\epsilon^k\} &= \exp\{A_{k+1}(\epsilon)\epsilon t\} + Q_k(\epsilon)\exp\{A_k(\epsilon)t\} - \sum_{p=0}^k Q_p(\epsilon) \\ &\quad \sum_{p=0}^{k-1} Q_p(\epsilon)\exp\{A_p(\epsilon)t/\epsilon^{k-p}\} \end{aligned} \quad (4.5.17)$$

$k=0,1,\dots,m$

Clearly,

$$\lim_{\epsilon \downarrow 0} \sup_{0 \leq t \leq T} \|\exp\{A_{k+1}(\epsilon)\epsilon t\} - I\| = 0 \quad (4.5.18)$$

Also, from Proposition 4.2.1,

$$\lim_{\epsilon \downarrow 0} \sup_{0 \leq t \leq T} \|\|Q_k(\epsilon)\exp\{A_k(\epsilon)t\} - Q_k\exp\{A_{k0}t\}\| = 0 \quad (4.5.19)$$

If $k \leq \ell$ in (4.5.7), then using (4.5.6), (4.5.8), (4.5.9) and (4.4.50) we get (4.5.3). To prove (4.5.4) notice that

$$\begin{aligned} P_\ell(\epsilon)\exp\{A_0(\epsilon)t/\epsilon^k\} &= P_\ell(\epsilon)\exp\{A_{k+1}(\epsilon)\epsilon t\} + Q_k(\epsilon)\exp\{A_k(\epsilon)t\} - \sum_{\substack{p=0 \\ p \neq \ell}}^k Q_p(\epsilon) \\ &\quad + \sum_{\substack{p=0 \\ p \neq \ell}}^{k-1} Q_p(\epsilon)\exp\{A_p(\epsilon)t/\epsilon^{k-p}\} \end{aligned}$$

$k=\ell+1,\dots,m$

and using again (4.5.6), (4.5.8), (4.5.9) and (4.4.50) we get:

$$\begin{aligned}
 P_\ell(\varepsilon) \exp\{A_o(\varepsilon)t/\varepsilon^k\} &= P_\ell + Q_k \exp\{A_{ko}t\} - \sum_{\substack{p=0 \\ p \neq \ell}}^k Q_p \\
 &= I + Q_k \exp\{A_{ko}t\} - \sum_{p=0}^k Q_p \\
 &= Q_k \exp\{A_{ko}t\} + P_o \dots P_k
 \end{aligned}$$

□

The above proposition differs from Proposition 4.5.2, proved for the US case, in the requirement that the exact expression for $P_\ell(\varepsilon)$ be used in (4.5.11) and not its zeroth order approximation, $P_\ell(0)$, as in (4.5.2). For Case 2 the problem is created by modes of the type

$$\varepsilon^\ell t^k e^{-\varepsilon^{\ell+p}} \sin \varepsilon^\ell t \tag{4.5.20}$$

At time scale t/ε^ℓ they look like growing modes $t^k \sin t$ and even though they are eventually attenuated, the time scale at which the attenuation takes place depends on the value of p . Furthermore, the maximum amplitude achieved before dying down goes to infinity as $\varepsilon \downarrow 0$.

By using the projection $P_\ell(\varepsilon)$ as in (4.5.11), we annihilate the term

$$Q_\ell(\varepsilon) \exp\{A_\ell(\varepsilon)\varepsilon^\ell t\} \tag{4.5.21}$$

in (4.4.29), which is the one responsible for unbounded amplitudes as $\varepsilon \downarrow 0$. It is possible to reduce these modes to $o(1)$ quantities

using only a finite number of terms in $P_\ell(\varepsilon) = P_\ell + \sum_{j=1}^{\infty} \varepsilon^j P_\ell^{(j)}$, but the number of terms required depends on the order of magnitude in ε of:

$$\sup_{t \leq 0} \left\| Q_\ell(\varepsilon) \exp\{A_\ell(\varepsilon) \varepsilon^\ell t\} \right\| \rightarrow \infty \quad \varepsilon \rightarrow 0 \quad (4.5.22)$$

and this, in turn, depends on the values of k and p in (4.5.20).

Thus, it is possible to substitute $P_\ell(\varepsilon)$ in (4.5.11) and (4.5.14) by

$$P_\ell^N(\varepsilon) = P_\ell + \sum_{j=1}^N \varepsilon^j P_\ell^{(j)}$$

provided N is such that

$$\sup_{t \geq 0} \left\| P_\ell^N(\varepsilon) \cdot Q_\ell(\varepsilon) \exp\{A_\ell(\varepsilon) \varepsilon^\ell t\} \right\| = o(1) \quad (4.5.23)$$

but the value of N depends on the specific structure of $A_\ell(\varepsilon)$. This concludes our discussion of Case 2.

Suppose now that $A_0(\varepsilon)$ is such that the sequence $A_{k,0}$ $k=0,1,\dots$, ends at $k=\ell$ because $A_{\ell,0}$ does not have SSNS, and that

$$\sum_{k=0}^{\ell} \text{rank } A_{k,0} < n \text{ rank } A_0(\varepsilon) \quad (4.5.24)$$

There is no problem in performing a time scale analysis up to time scale t/ε^ℓ , as state in Proposition 4.5.3. Equation (4.5.24) indicates, however, that there are eigenvalues of $A_0(\varepsilon)$ that of order $o(\varepsilon^\ell)$ whose

effect will not be captured by this partial time scale analysis. On the other hand, although the limit of $\exp\{A_0(\epsilon)t/\epsilon^{\ell+1}\}$ as $\epsilon \downarrow 0$ cannot exist because $\|\exp\{A_0(\epsilon)t/\epsilon^{\ell+1}\}\| \rightarrow \infty$, it may well be possible that the system has well defined behavior at slower time scales, when the effects of high amplitude transients have dissappeared. The techniques we have used, however, do not seem to be adequate to treat this case.

The asymptotic analysis of singularly perturbed systems that violate the MSST condition remains largely an open question. See, however, Chapter VI for an indication of how this analysis could be carried out, and the work of [Ver 81] and [Sas 80] for related results from a singular systems and root-loci point of view respectively.

4.6 Summary and Conclusions

In this chapter we have studied the asymptotic behavior of $\exp\{A_0(\epsilon)t\}$ over the infinite time interval $[0, \infty)$ for

$$A_0(\epsilon) = \sum_{k=0}^{\infty} \epsilon^k A_{0k}$$

We have formalized the notion of multiple time scale behavior and that of reduced order models of the system

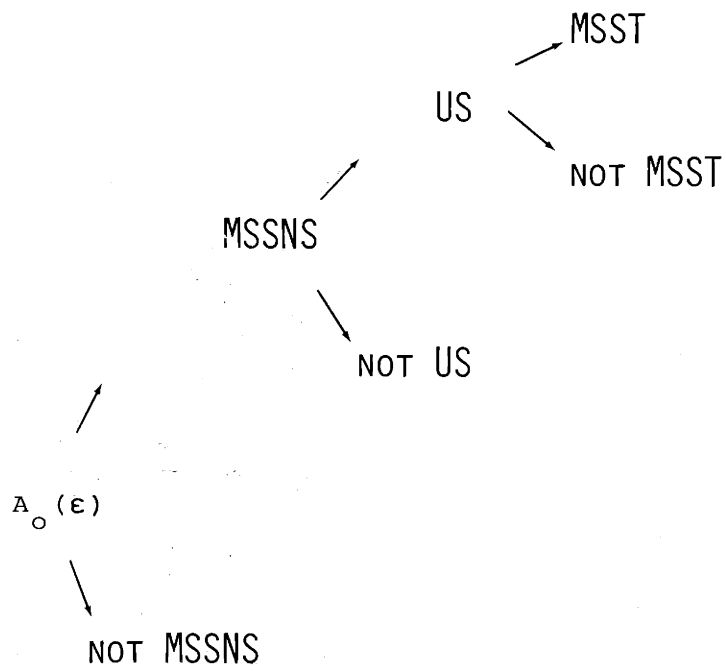
$$\frac{dx(t)}{dt} = A_0(\epsilon)x(t)$$

valid at different time scales, and we have identified several conditions on $A_0(\epsilon)$ which give rise to qualitatively different asymptotic behavior of (5.19). The hierarchical relationship among these conditions is visualized in Fig. 4.3.

The most important result is probably the fact that multiple semistability is a necessary and sufficient condition for a system to have well defined behavior at all its time scales, and that the fundamental time scales of such systems are of the form t/ϵ^k , with k integer. If a system does not satisfy the multiple semistability condition then a time scale analysis of it will not be adequate to capture completely all the features of the system's behavior. Conversely, from a modelling viewpoint, if a system has a well defined multiple time scale behavior, then proper modelling must result in a system matrix that satisfies the MSST condition.

For MSST systems we have developed a methodology for the computation of the different reduced-order models of a system which describe its evolution at different time scales, and we have shown that these-reduced-order models can be combined to approximate the original system.

In the next chapter we show that for an important class of systems, namely, those modelled by Finite State Markov Processes (or any other positive uniformly stable system), the MSST condition



MSSNS = Multiple Semisimple Nullstructure
US = Uniform Stability
MSST = Multiple Semistability

Figure 4.3: Different cases studied in chapter 4.

is always satisfied indicating that the results developed in this chapter have indeed a wide range of applicability. Furthermore, in these cases we can give important and useful interpretations of the hierarchy of reduced-order models in terms of the aggregation of the processes they represent.

CHAPTER V: HIERARCHICAL AGGREGATION OF FINITE
STATE MARKOV PROCESSES

5.1 Introduction and Overview

In this chapter we apply the results developed in Chapter IV to singularly perturbed Finite State Markov Processes (FSMP's). We show that for this important class of models the complete time scale decomposition of developed in Chapter IV can *always* be performed. Furthermore, as we will see, the reduced-order models introduced in Section 4.4.3 have in this setting a strong intuitive appeal forming a hierarchy of aggregated models.

After a review of other author's work in this area and a motivating example, the bulk of the results is presented in Sections 5.4 and 5.5. The asymptotic approximations and the multiple time scale behavior introduced in Chapter IV lead to the consideration of stochastically discontinuous Markov processes which in the past have received little attention. In Section 5.4, such processes are studied and we interpret their properties as the limiting behavior of FSMP's with rare events.

5.2 Survey of Related Work

We discuss here in more detail the work of Delebecque [Del 82] mentioned in Chapter II. This work addresses the same problem we consider in this chapter, i.e., the asymptotic behavior of a FSMP with generator

$$A_o(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p A_{op} \quad (5.2.1)$$

Delebecque deals mainly with the discrete time case while we consider the continuous time case, but this does not prevent a comparison of approaches and results.

There are two basic methodologic differences between our approach and that in [Del 82]. We do all our analysis in the time domain using the results in Chapter IV that we have derived using Kato's perturbation theory. Delebecque, on the contrary, works in the frequency domain (i.e., dealing with convergence of expressions involving the resolvent matrix) and it states several results obtained by *formally* equating terms in series in powers of ϵ .

Starting from (6.2.1) he formally introduces a "reduced" series

$$\bar{A}_1(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p \bar{A}_{1p} \quad (5.2.2)$$

as follows:

$$\begin{aligned} \bar{A}_{10} &= P_0 A_{01} P_0 \\ \bar{A}_{11} &= P_0 A_{02} P_0 - P_0 A_{01} A_{00}^{\#} A_{01} P_0 \\ &\vdots \\ &\vdots \\ \bar{A}_{1p} &= \sum_{r=1}^{p+1} \sum_{k_1 + \dots + k_r = n+1} P_0 A_{0, k_1} A_{00}^{\#} \dots A_{00}^{\#} A_{0, k_r} P_0 \end{aligned}$$

where P_0 is the ergodic projection of the unperturbed chain $A_0(0) = A_{00}$, and he shows that $\bar{A}_{10} = P_0 A_{01} P_0$ can be interpreted as the generator of a Markov chain in the A_{00} -invariant subspace $R(P_0)$ which has one state per each ergodic class of A_{00} . Repeating now the same construction with the series $\bar{A}_1(\epsilon)$ he arrives at a new chain $\bar{A}_{20} = P_1 \bar{A}_{11} P_1$ defined now on range of P_1 , the ergodic projection of the previous chain \bar{A}_{10} and the recursion can go on up to a certain stage \bar{A}_{m0} .

This construction is very similar to the recursion process we have discussed in Section 4.4 and that we apply in this chapter to analyze FSMP's. Both schemes result in the same matrices $\bar{A}_{10}, \bar{A}_{20}, \dots, \bar{A}_{m0}$ which are proven to be Markov generators although Delebecque's series probably involves less computations (we have already discussed this point in Section 4.4.5).

Delebecque's main result ([Del 82], Theorem 1) is that the matrices $\bar{A}_{10}, \bar{A}_{20}, \dots, \bar{A}_{m0}$ define a sequence of Markov generators each with a number of states equal to the number of ergodic classes of the preceding chain. A result that our analysis in this chapter confirms.

Other claims made in [Del 82] do not seem to be as well substantiated. For example, it is claimed that $\bar{A}_1(\epsilon)$ itself is a Markov generator for $\forall \epsilon \geq 0$ but the proof given seems to imply only that $\bar{A}_{10} + \epsilon \bar{A}_{11}$ is such a generator. Furthermore, the interpretation of

the sequence of generators \bar{A}_{ko} as limiting aggregated models is in some senses incomplete and does not expose the basic structure and concepts in a clear fashion. It is claimed that if $\eta^\varepsilon(t)$ is the chain with generator $A_o(\varepsilon)$ then $\eta^\varepsilon(t/\varepsilon^k)$ converges weakly to a FSMP with generator \bar{A}_{ko} . This is not possible since the limit process has different number of states than the chain with generator \bar{A}_{ko} . An aggregation operation is needed to collapse several states of $\eta^\varepsilon(t/\varepsilon^k)$ into a single state before interpreting the limit results. In this respect, Delebecquec's work lacks the notion of stochastic discontinuity that we use in Section 5.4 to introduce the need for aggregation.

The results presented in this chapter include those in [Del 82] and go beyond in several respects:

- i) We give a precise interpretation of the aggregated chains as the limiting behavior of aggregated versions of the original process $\eta^\varepsilon(t)$; and, more importantly,
- ii) we show that these aggregated models can then be combined to produce an asymptotic approximation of the original process which is uniformly valid over $t \in [0, \infty)$.
- iii) From a technical viewpoint our derivations are not formal as in [Del 82] but one rigorously based on Kato's perturbation theory of linear operators and they are seen to be a particular case of a

more general aggregation methodology for linear systems. Our methodology is more powerful and leads to precise statements which provide a very clear and complete picture of the multiple time scale structure not found in any other work.

iv) Our analysis in the time domain leads to the consideration of stochastically discontinuous processes which aside from being interesting in their own, are found to play an important role in approximations involving multiple time scale behavior.

5.3 A Motivating Example

Consider the process $\eta^\varepsilon(t)$ portrayed in Figure 5.1 and suppose that $\eta^\varepsilon(0) \in \{e_1, e_2\}$. For $\varepsilon > 0$, the process will spend a random amount of time switching between e_1 and e_2 and eventually it will get trapped in e_3 . It is clear that we can identify phenomena occurring at two time scales. At the "fast" time scale only transitions between e_1 and e_2 occur and $\eta^0(t)$ (in which there is no possibility of transition to e_3) is a good model for that. At the "slow" time scale the important phenomena is a transition to state e_3 . Suppose we are interested only in the phenomena occurring at the slow time scale and for ε very small. It is then logical to study the process $\eta^\varepsilon(t/\varepsilon)$ in the limit as $\varepsilon \downarrow 0$.

Figure 5.2 shows a typical sample function of $\eta^\varepsilon(t/\varepsilon)$. Each sojourn in states e_1 and e_2 has an average duration of order ε , and on the order of $1/\varepsilon$ such sojourns take place before absorption. In the limit as $\varepsilon \downarrow 0$ the sample functions of $\eta^\varepsilon(t/\varepsilon)$ approach functions with an infinite number of discontinuities on finite time intervals. In fact, as we will see in Section 5.5, the finite dimensional distributions of $\eta^\varepsilon(t/\varepsilon)$ converge to those of a stochastically discontinuous Markov process $\eta_1(t)$ with sample functions of the type shown in Figure 5.3. Furthermore, the time to absorption, τ^ε , is the sum of a geometrically distributed number of i.i.d. positive random variable and it has mean of order 1. Using results in [Kei 78] we can conclude that it converges to an exponentially distributed

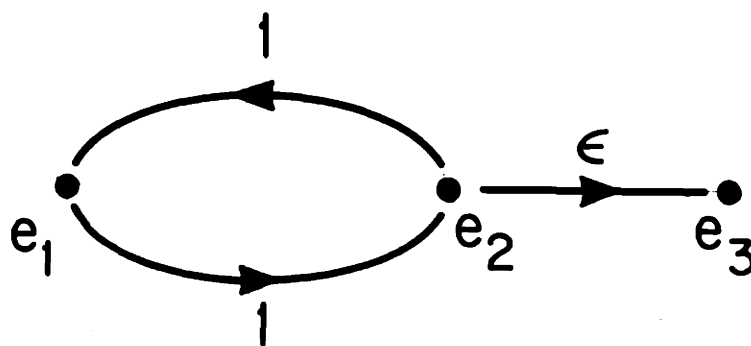


Figure 5.1: The process $\eta^\epsilon(t)$.

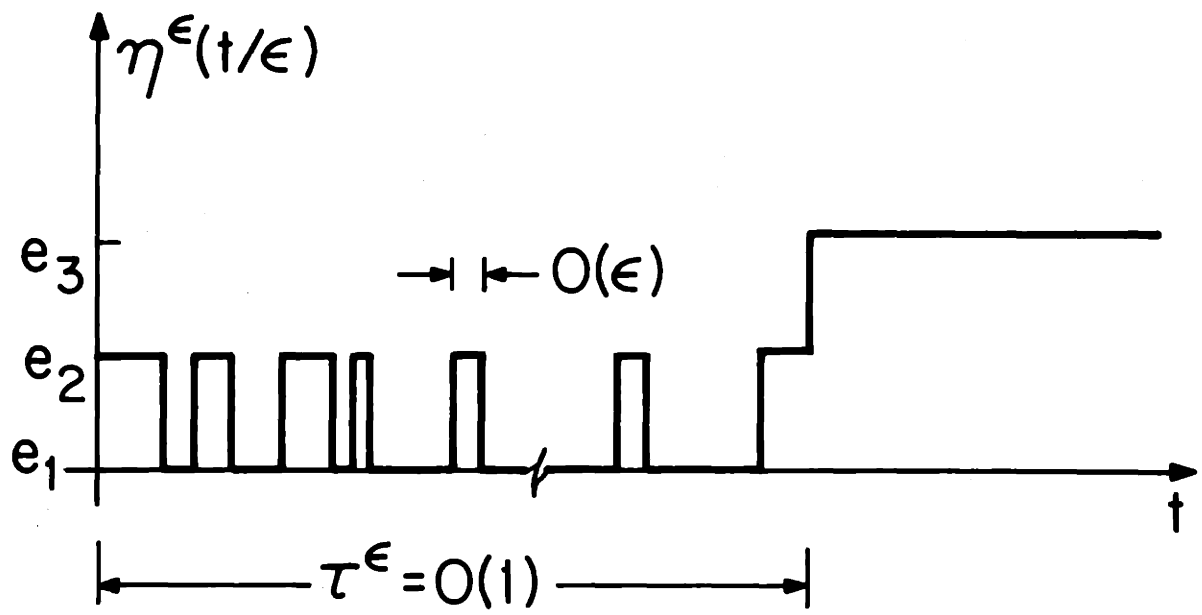


Figure 5.2: A typical sample function of $\eta^\epsilon(t/\epsilon)$

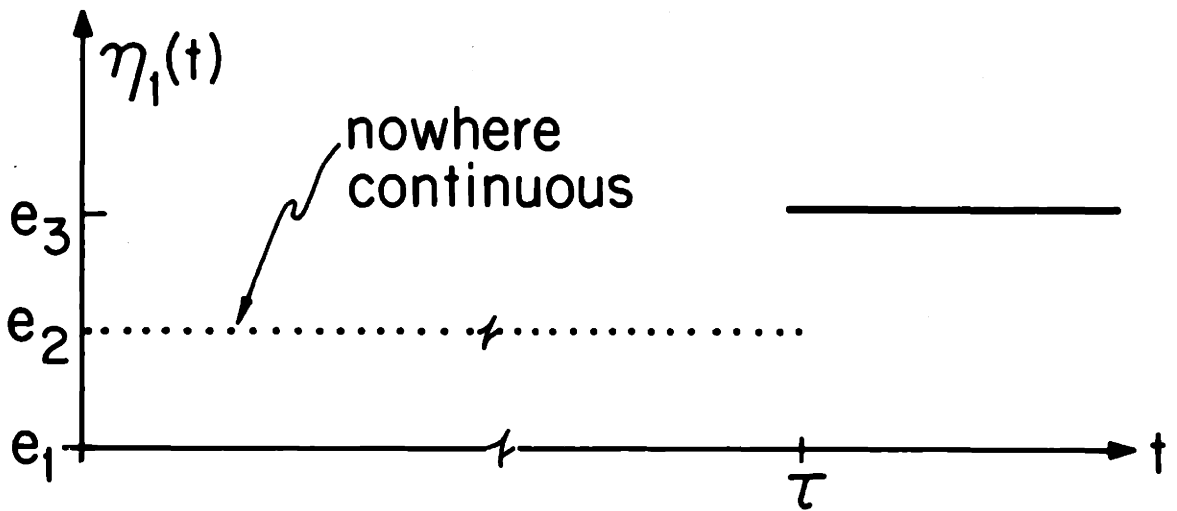


Figure 5.3: A sample function of the limiting process

$$\eta_1(t) \triangleq \lim_{\epsilon \downarrow 0} \eta^\epsilon(t/\epsilon).$$

random variable τ . Therefore, if we define a new process $\hat{\eta}_1(t)$ by

$$\hat{\eta}_1(t) = \begin{cases} \hat{e}_1 & \text{if } \eta_1(t) \in \{e_1, e_2\} \\ \hat{e}_2 & \text{if } \eta_1(t) = e_3 \end{cases}$$

it is clear that $\hat{\eta}_1(t)$ is the Markov process shown in Figure 5.4 which can be thought of as an approximate, aggregated model for the slow behavior of $\eta^\epsilon(t)$.

This example indicates the need to deal with stochastically discontinuous processes when analyzing the multiple time scale behavior of singularly perturbed FSMP's. Stochastic discontinuity reflects the fact that when a specific time scale is selected, each transition that is likely to occur at a faster time scale (if any) appears to occur instantaneously upon entering some state. Aggregation in this context is simply the avoidance of stochastic discontinuity by discarding the details modeled by these faster, asymptotically discontinuous transitions.

The rest of this chapter is devoted to making these ideas precise, and to show that they generalize to arbitrary FSMP's with generator of the form

$$A_o(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p A_{op} \quad (5.3.1)$$

and phenomena occurring at several different time scales.

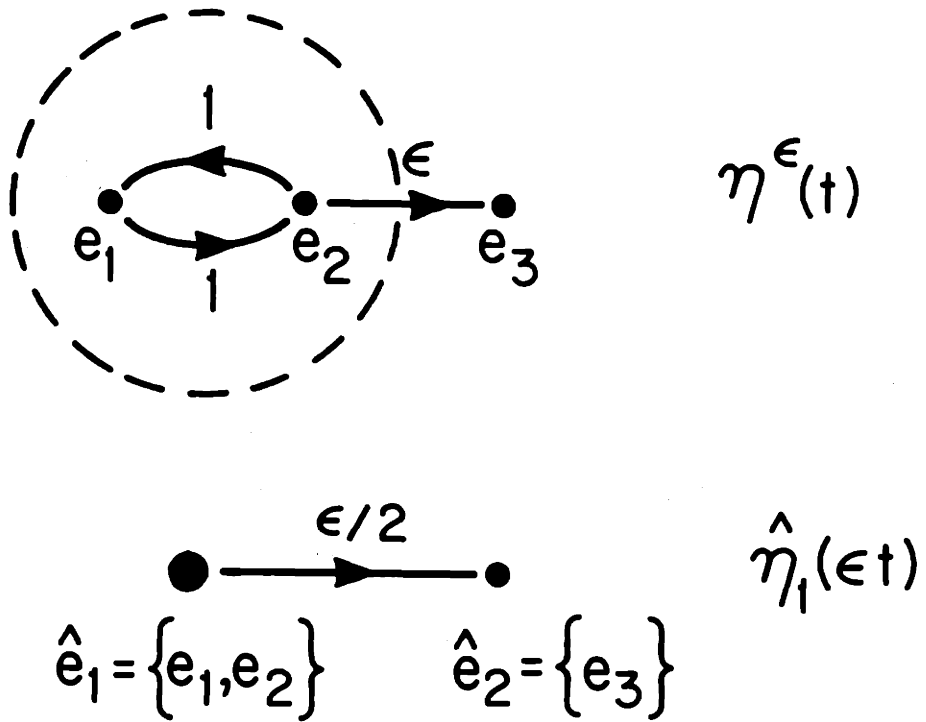


Figure 5.4: The aggregated model $\hat{\eta}_1(\epsilon t)$.

5.4 Stochastically Discontinuous FSMP's

We consider here continuous-time, stationary, finite-state Markov processes $\{\eta(t), t \geq 0\}$ that may undergo an infinite number of transitions in finite time intervals. Such processes violate the continuity condition:

$$\lim_{t \downarrow 0} \Pr\{\eta(t) = \eta(0)\} = 1$$

and, accordingly, are referred to as *stochastically discontinuous* [Dyn 65]. They were first analyzed in [Doe 38] and [Doo 42] but were considered pathological from an applications view point and since then stochastic continuity has been a standard assumption in the literature (see for example [Sko 65] and [Wil 79]). As we have indicated in Section 5.3, however, stochastically discontinuous processes are obtained as limits of Markov processes with transition rates of different orders of magnitude and the stochastic discontinuity property has a natural and important interpretation in this context.

In this section we carry out an analysis of FSMP's along the same lines usually followed for stochastically continuous processes (as in [Doo 53], for example), but for the general, stochastically discontinuous case.

A stationary Markov process $\{\eta(t), t \geq 0\}$ taking values in a finite state space $E = \{e_1, e_2, \dots, e_n\}$ is completely described by its transition

probability matrix $P(t)$ whose elements are the transition probabilities:

$$p_{ij}(t) = \Pr\{\eta(t) = j \mid \eta(0)=i\} \quad i, j \in E, t \geq 0$$

An $(n \times n)$ matrix-valued function $P(t)$, $t \geq 0$ is a transition probability matrix of some FSMP if and only if it satisfies the following conditions:

$$i) \quad P(0) = I \quad (5.4.3)$$

$$ii) \quad P(t) \geq 0, \quad \forall t \geq 0 \quad (5.4.4)$$

$$iii) \quad P(t) \cdot \quad = \quad * \quad (5.4.5)$$

$$iv) \quad P(t)P(\tau) = P(t+\tau), \quad \forall t, \tau \geq 0 \quad (5.4.6)$$

In addition, it is known (see [Doe 38], [Doo 42]) that if $P(t)$ is the transition probability matrix of a FSMP then it is continuous for $t > 0$ and the limit

$$\lim_{t \downarrow 0} P(t) = \Pi \quad (5.4.7)$$

always exists. It follows from (5.4.4)-(5.4.6) and the continuity of $P(t)$ that Π satisfies:

$$\Pi > 0, \quad \Pi \cdot \quad = \quad , \quad \Pi^2 = \Pi \quad (5.4.8)$$

and also

$$\Pi P(t) = P(t) \Pi = P(t) \quad (5.4.9)$$

* $\mathbb{1} = [1, 1, \dots, 1]$.

If Π is the identity matrix then the process $\eta(t)$ with transition probability matrix $P(t)$ is called *stochastically continuous* otherwise it is called *stochastically discontinuous*.

Theorem 5.4.1

If $P(t)$ is the transition probability matrix of a FSMP then,

$$P(t) = \Pi \exp\{At\} \quad t > 0 \quad (5.4.10)$$

for a pair of matrices Π, A satisfying:

$$(i) \quad \Pi \geq 0, \Pi \cdot \quad (5.4.11)$$

$$(ii) \quad \Pi A = A \Pi = A ; \quad (5.4.12)$$

$$(iii) \quad A \cdot = 0 ; \quad (5.4.13)$$

$$(iv) \quad A + c\Pi \geq 0 \quad \text{for some } c > 0 . \quad (5.4.14)$$

Conversely, any pair of matrices A, Π satisfying

(i)-(iv) uniquely determine a FSMP with transition probability matrix given by (5.4.10).

Proof: The proof of (5.4.10) given here adapts a more general result on semigroups in [Hil 53] to the context of FSMP's. By the continuity properties of $P(t)$ we have:

$$\lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} P(\tau) d\tau = \begin{cases} P(t) & \text{if } t > 0 \\ \Pi & \text{if } t = 0 \end{cases}$$

for some Π satisfying (5.4.8) and (5.4.9).

It follows from (5.4.6) and (5.4.9) that $\forall t > 0$,

$$\int_0^t P(h+\tau) d\tau - \int_0^t P(\tau) d\tau = (P(h) - \Pi) \int_0^t P(\tau) d\tau$$

which gives

$$\frac{1}{h} \int_t^{t+h} P(\tau) d\tau - \frac{1}{h} \int_0^h P(\tau) d\tau = \frac{1}{h} (P(h) - \Pi) \int_0^t P(\tau) d\tau \quad (5.4.15)$$

As $h \downarrow 0$ the left-hand side converges to $P(t) - \Pi$ and therefore

$$\lim_{h \downarrow 0} \frac{P(h) - \Pi}{h} \triangleq A \quad (5.4.16)$$

exists. Taking limits as $h \downarrow 0$ in (5.4.15) we get

$$P(t) = \Pi + A \int_0^t P(\tau) d\tau \quad \forall t > 0$$

establishing (5.4.10). Definition (5.4.16) together with (5.4.8)

and (5.4.9) give (5.4.12), and (5.4.13) follows immediately from

(5.4.10) and the fact that $\Pi \cdot = \cdot$. The positivity of $P(t)$, i.e.,

$$\frac{1}{h} P(h) = \frac{1}{h} \Pi \exp\{At\} = \frac{1}{h} \Pi + A + \frac{o(h)}{h} \geq 0 \quad \text{for } h \geq 0$$

implies that for h small enough $A + \Pi/h \geq 0$ establishing (5.4.14).

To prove the converse suppose now that Π and A satisfy (5.4.11)-

(5.4.14). Then, $P(t) = \Pi \exp\{At\}$ clearly satisfies (5.4.5) and (5.4.6)

and the positivity condition follows from (5.4.14) as indicated below:

$$\begin{aligned} \Pi \exp\{At\} &= \Pi e^{-ct} \exp\{(A + cI)t\} \\ &= e^{-ct} \Pi \sum_{n=0}^{\infty} \frac{(A\Pi + c\Pi)^n}{n!} t^n \geq 0 \end{aligned}$$

We shall refer to the projection $\Pi = \lim_{t \downarrow 0} P(t)$ as the *ergodic projection at zero* and to the matrix

$$A = \lim_{h \downarrow 0} \frac{P(h) - \Pi}{h} \tag{5.4.17}$$

as the *infinitesimal generator* of $P(t)$. □

Remark:

1) It follows from (5.4.11) that Π is the matrix of ergodic probabilities of a Markov chain and as such it determines a partition of E in terms of ergodic classes, E_i° , $i=1, \dots, s$, and transient states,

E_T° ,

$$E = \left(\bigcup_{i=1}^s E_i^\circ \right) \cup E_T^\circ$$

that we will refer to as the *ergodic partition at zero*. As we will see later, this partition corresponds to a classification of states into different types. While the process is in absorbing state (i.e. in an ergodic class E_i^0 with a single element), the process behaves as a stochastically continuous FSMP. Instantaneous transitions occur between states belonging to the same ergodic class at zero, and transient states are visited only during transitions between ergodic classes, with no time spent in them.

2) For stochastically continuous processes $\Pi = I$ and conditions (i) - (iv) only require that the rows of A add up to zero and that all its off-diagonal entries be non-negative. In the general case some off-diagonal entries of A can be negative provided the corresponding entry in Π is non-zero (see Example 5.4.2 below). The usual interpretation of a_{ij} as the rate of transitions from state i to j is thus no longer valid in the stochastically discontinuous case. To interpret these entries it is first necessary to perform an aggregation as discussed later in this section.

Example 5.4.2

The following is a stochastically discontinuous transition probability matrix:

$$P(t) = \begin{bmatrix} p_1 e^{-\lambda t} & p_2 e^{-\lambda t} & 1 - e^{-\lambda t} \\ p_1 e^{-\lambda t} & p_2 e^{-\lambda t} & 1 - e^{-\lambda t} \\ 0 & 0 & 0 \end{bmatrix}, \quad p_1 + p_2 = 1, \quad t > 0$$

with initial projection and infinitesimal generator given by:

$$\Pi = \begin{bmatrix} p_1 & p_2 & 0 \\ p_1 & p_2 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad A = \begin{bmatrix} -p_1 \lambda & -p_2 \lambda & \lambda \\ -p_1 \lambda & -p_2 \lambda & \lambda \\ 0 & 0 & 0 \end{bmatrix}$$

For $p_1 = p_2 = \lambda = 1/2$ this is the stochastically discontinuous limit process $\eta_1(t)$ described in Section 5.4.1.

5.4.1 Implications of Stochastic Discontinuity

If we consider a separable version of a stochastically continuous FSMP then its sample functions are easily visualized as piecewise continuous functions taking values in E [Doo 53]. The evolution of the process can be thought of as succession of stays in different states of E , each being of random duration and exponentially distributed.

The sequence of states visited follows a Markov chain law with one-step transition probabilities determined by the entries of the generator A . On the contrary, the sample functions of a stochastically discontinuous process are much more irregular. As we will now prove, these processes have instantaneous states, i.e., states in which the process spends no time with probability one. Furthermore, in general, a stochastically discontinuous process spends a non-zero amount of time switching among instantaneous states. The sample functions are therefore nowhere continuous on certain time intervals.

Consider a separable version of A FSMP $\eta(t)$ with initial projection Π and generator A , and let Λ be a separating set. For $t > 0$ and $n=0,1,\dots$, take

$$0 = t_{0n} < t_{1n} < \dots < t_{nn} = t$$

in such a way that the sets

$$\Lambda_n = \{t_{0n}, t_{1n}, \dots, t_{nn}\}$$

increase monotonically and $\cup \Lambda_n = \Lambda \cap [0, t]$. Then we have:

$$\Pr\{\eta(\tau) = i, \forall \tau \in [0, t] \mid \eta(0) = i\} =$$

$$\Pr\{\eta(\tau) = i, \forall \tau \in [0, t] \cap \Lambda \mid \eta(0) = i\} =$$

$$\lim_{n \rightarrow \infty} \Pr\{\eta(\tau) = i, \forall \tau \in [0, t] \cap \Lambda_n \mid \eta(0) = i\} =$$

$$\lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} P_{ii}(t_{k+1,n} - t_{k,n}) =$$

$$\exp\left\{ \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} \log P_{ii}(t_{k+1,n} - t_{k,n}) \right\} \quad (5.4.18)$$

where $P_{ii}(t)$ are the diagonal elements of $P(t) = \Pi \exp\{At\}$.

Equation (5.4.18) facilitates a classification of the states of $\eta(t)$ according to the diagonal entries, π_{ii} , of Π . If $\pi_{ii} = 0$ then $P_{ii}(h) \rightarrow 0$ as $h \rightarrow 0$ and therefore (5.4.18) gives:

$$\Pr\{\eta(\tau) = i, \forall \tau \in [0, t] \mid \eta(0) = i\} = 0 \quad \forall t > 0$$

If, on the other hand, $0 < \pi_{ii} \leq 1$ use (5.4.17) to write:

$$\frac{P_{ii}(h)}{h} = 1 + \frac{a_{ii}}{\pi_{ii}} h + o(h)$$

or

$$\log P_{ii}(h) = \log \pi_{ii} + \frac{a_{ii}}{\pi_{ii}} h + o(h)$$

and it follows from (5.4.18) that

$$\Pr\{\eta(\tau) = i, \forall \tau \in [0, t] \mid \eta(0) = i\} = \begin{cases} 0 & \text{if } \pi_{ii} < 1 \\ \exp\{a_{ii}t\} & \text{if } \pi_{ii} = 1 \end{cases} \quad (5.4.19)$$

Definition 5.4.3

A state i will be called *instantaneous* if $\pi_{ii} < 1$ and *regular* if $\pi_{ii} = 1$. An instantaneous state j will be called *evanescent* if $\pi_{jj} = 0$.

Remarks

- (1) Notice that this classification is based on the ergodic partition at zero.
- (2) We have just seen that the sojourn time in instantaneous states is zero w.p.1.
- (3) Also, the sojourn time in regular states is exponentially distributed. All states of a stochastically continuous process are regular.
- (4) In Example 5.4.2, states {1,2} are instantaneous, non-evanescent states while 3 is regular.
- (5) Even though the duration of stays in a given instantaneous state is zero w.p.1, there is, in general, a non-zero probability of finding the process in an instantaneous state at any given time (as in states {1,2} of Example 5.4.2).
- (6) The probability of finding the process in an evanescent state at any given time is zero. This follows from the fact that $\pi_{ii} = 0$ implies $\pi_{ji} = 0 \quad j=1, \dots, n$ (i.e., evanescent states are transient states of the chain Π) and because:

$$P(t) = \Pi \exp\{At\} = \Pi \exp\{At\} \quad \forall t > 0$$

we have

$$p_{ji}(t) = 0, \quad \forall t > 0, \quad j=1, \dots, n .$$

The evanescent states can thus be neglected in the sense that there exists a version of the process $\eta(t)$ with the same finite dimensional distributions which does not take values in the set of evanescent states.

As we will see, the evolution of a stochastically continuous FSMP can be thought of as follows: While in a regular state, it behaves as a stochastically continuous process. Upon entering a state belonging to, say, E_k^o , for some $s+1 \leq k \leq r$, the process starts switching instantaneously among states in E_k^o . The amount of time spent in E_k^o is exponentially distributed and after a random stay in E_k^o the process jumps to some state in $E - E_k^o$. Evanescent states may be visited during transitions between ergodic classes but, as we said, they can be pruned without affecting the finite dimensional distributions of $\eta(t)$.

5.4.2 Aggregation of Stochastically Discontinuous FSMP's

We prove now that all probabilistic properties of a stochastically discontinuous process can be derived from its ergodic projection at zero plus an aggregated version of the process that is stochastically continuous.

The following well known proposition establishes notation that we use in the sequel.

Proposition 5.4.4

Let Π be the ergodic projection at zero of a FSMP

then, by an adequate ordering,

$$\Pi = \begin{bmatrix} \Pi_{11} & 0 & \dots & 0 \\ 0 & \Pi_{22} & \dots & 0 \\ \vdots & & \Pi_{ss} & 0 \\ 0 & & & \\ \Pi_{1,s+1} & \dots & \Pi_{s,s+1} & 0 \end{bmatrix} \quad (5.4.20)$$

with $\Pi_{kk} = \mu_k \mu_k^T$, $k=1, \dots, s$, for some vector $\mu_k > 0$ such that $\mu_k^T \cdot =$; and $\Pi_{k,s+1} = \delta_k \mu_k^T$, $k=1, \dots, s$ for a set of vectors $\delta_k > 0$ such that $\sum_{k=1}^s \delta_k =$.

Furthermore, define the $(n \times s)$ matrix V and the $(s \times n)$ matrix U as follows:

$$V = \begin{bmatrix} 0 & 0 & & 0 \\ 0 & 0 & \dots & \\ \delta_1 & \delta_2 & \dots & \delta_s \end{bmatrix} \quad U = \begin{bmatrix} \mu_1^T & 0 & \dots & 0 & 0 \\ 0 & \mu_2^T & \dots & 0 & 0 \\ 0 & 0 & \dots & \mu_s^T & 0 \end{bmatrix} \quad (5.4.21)$$

then,

$$\begin{aligned} V \cdot U &= \Pi \\ U \cdot V &= I \end{aligned} \quad (5.4.22)$$

Proof: See [Doo 42]. The vector μ_k is the vector of ergodic probabilities of a Markov chain with state space E_k^o and transition matrix Π_{kk} . The vectors δ_k are the trapping probabilities from transient states to ergodic classes. □

Remark: Notice that the structure of (5.4.20) makes explicit the ergodic partition at zero.

We shall refer to (5.4.22) as the canonical product decomposition of Π . Notice that U and V satisfy

$$U \cdot \quad = \quad (5.4.23)$$

$$V \cdot \quad = \quad (5.4.24)$$

$$U \cdot \Pi = U \quad (5.4.25)$$

$$\Pi \cdot V = V \quad (5.4.26)$$

Theorem 5.4.5

Let $P(t) = \Pi \exp\{At\}$ be the transition probability matrix of a FSMP $\eta(t)$ taking values in $E = \{e_1, \dots, e_n\}$ and let s be the number of ergodic classes at zero. Let $\Pi = V \cdot U$ be the canonical product decomposition of Π . Then:

$$\hat{P}(t) \triangleq U P(t) V = \exp\{UAVt\} \quad \forall t > 0 \quad (5.4.27)$$

is the transition probability matrix of a FSMP taking values in $\hat{E} = \{\hat{e}_1, \dots, \hat{e}_s\}$ and

$$P(t) = V \hat{P}(t) U \quad \forall t > 0 \quad (5.4.28)$$

Proof: $\hat{P}(t) \geq 0$, $\hat{P}(t) \cdot \quad = \quad$ and $\hat{P}(t)\hat{P}(t) = \hat{P}(t+\tau)$ follow immediately from positivity of U and V , from (5.4.23) and (5.4.24), and from (5.4.9) and (5.4.22), respectively. Use now $I = U\Pi V$ and $A = VUA$ to write

$$\begin{aligned}\hat{P}(t) &= U\Pi V + \sum_{k=1}^{\infty} \frac{t^k}{k!} UA (VUA)^{k-1}V \\ &= I + \sum_{k=1}^{\infty} \frac{t^k}{k!} (UAV)^k = \exp\{UAV t\}\end{aligned}\tag{5.4.29}$$

To prove (5.4.28) notice that

$$V\hat{P}(t)U = VUP(t)VU = \Pi P(t)\Pi = P(t)\tag{5.4.30}$$

□

Equation (5.4.27) can be interpreted as performing an aggregation operation that masks the stochastically discontinuous nature of $P(t)$. Define the aggregated process $\hat{\eta}(t)$ taking values in $\hat{E} = \{\hat{e}_1, \dots, \hat{e}_s\}$ as follows:

$$\hat{\eta}(t) = e_i \quad \text{if } \eta(t) \in E_i^o \quad i=1, \dots, s\tag{5.4.31}$$

Assuming that we deal with a version of $\eta(t)$ which does not take values in E_T^o , $\hat{\eta}(t)$ is well defined for $t \geq 0$.

Corollary 5.4.6

The aggregated process $\hat{\eta}(t)$ is a stochastically continuous FSMP with transition probability matrix $\hat{P}(t)$, i.e.,

$$\Pr\{\eta(t) \in E_j^o | \eta(0) = e_k \in E_i^o\} = \Pr\{\hat{\eta}(t) = \hat{e}_j | \hat{\eta}(0) = \hat{e}_i\} = \hat{p}_{ij}(t) \quad \forall t > 0$$

Proof: Follows directly from (5.4.31) and the structure of the matrices U and V .

□

Note also that (5.4.28) can be interpreted as follows:

$$\Pr\{\eta(t) = e_i | \eta(0) = e_j\} = \mu_\ell^i \cdot \Pr\{\hat{\eta}(t) = \hat{e}_\ell | \hat{\eta}(0) = \hat{e}_p\} \quad (5.4.32)$$

$$e_j \in E_p^\circ, e_i \in E_\ell^\circ$$

where μ_ℓ^i is the i^{th} component of the ergodic probability vector μ_ℓ . That is, the transitions between the ergodic classes E_i° are governed by the aggregated process while, once in one of the classes E_i° , the ergodic probabilities μ_ℓ^i are immediately established due to the instantaneous nature of the transitions.

Example 5.4.7

Consider the process $\eta(t)$ in Example 5.4.2. Its ergodic partition at zero is

$$E_1^\circ = \{3\} ; E_2^\circ = \{1,2\}$$

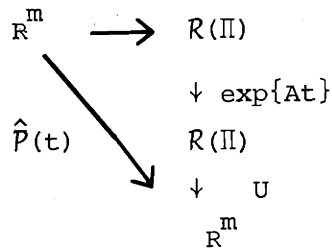
Its aggregated version $\hat{\eta}(t)$ has two states one of which corresponds to the consolidation of states 1 and 2 of $\eta(t)$. The canonical product decomposition of Π is

$$\Pi = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} p_1 & p_2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and matrix of transition rates for the aggregated process $\hat{\eta}(t)$ is

$$A = U \cdot \begin{bmatrix} -p_1\lambda & -p_2\lambda & \lambda \\ -p_1\lambda & -p_2\lambda & \lambda \\ 0 & 0 & 0 \end{bmatrix} \cdot V = \begin{bmatrix} -\lambda & \lambda \\ 0 & 0 \end{bmatrix}$$

In view of the interpretations described above, we shall refer to (5.4.27) as the *aggregation operation* and to (5.4.28) as the *disaggregation operation*. These operations can also be interpreted from a geometrical point of view. Notice that the stochastically discontinuous transition probability matrix $P(t) = \Pi \exp\{At\}$ defines a transition matrix on $R(\Pi)$ which is continuous at zero. By construction, the matrix V maps R^S into $R(\Pi)$ on a one-to-one basis and U maps $R(\Pi)$ back into R^S also one-to-one. We thus have the following diagram:



From this point of view the aggregation operation is interpreted as a restriction to the range of Π of the domain of definition of the generator A .

To conclude this section notice that all relevant information about a stochastically discontinuous process is contained in its aggregated version and its ergodic projection at zero. Therefore, the analysis of such processes can be reduced, using Theorem 5.4.5, to the well known stochastically continuous case. As an example of reduction consider the following corollary of Theorem 5.4.5.

Corollary 5.4.8

If $P(t) = \Pi \exp\{At\}$ is the transition probability matrix of a FSMP then the limit

$$\lim_{t \rightarrow \infty} P(t) \stackrel{\Delta}{=} \Pi^{\infty} \tag{5.4.33}$$

always exists and satisfies

- i) $\Pi^{\infty} \geq 0$, $\Pi^{\infty} \cdot \quad = \quad , (\Pi^{\infty})^2 = \Pi^{\infty}$
- ii) $\Pi^{\infty} P(t) = P(t) \Pi^{\infty} = \Pi^{\infty}$
- iii) $A\Pi^{\infty} = \Pi^{\infty}A = 0$
- iv) $\Pi^{\infty}\Pi = \Pi \cdot \Pi^{\infty} = \Pi^{\infty}$

Proof: Follows from (5.4.28) and the fact that (5.4.33) always exists for stochastically continuous processes [Doo 53]. □

In the sequel we shall refer to Π^{∞} as the *ergodic projection at ∞* .

For future reference it is important to notice that because

$$P(t) = \Pi \exp\{At\} = \exp\{At\} - I + \Pi$$

equation (5.4.33) implies that generators of FSMP's are semistable matrices.

5.5 Singularly Perturbed FSMP's and Aggregated Models

Consider now a stochastically continuous FSMP $\eta^\epsilon(t)$ that takes values in $E_o = \{e_1, \dots, e_{n_o}\}$ with infinitesimal generator of the form:

$$A_o(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p A_{op} \quad \epsilon \in [0, \epsilon_o] \quad (5.5.1)$$

The small parameter ϵ models rare transitions in $\eta^\epsilon(t)$ and we shall refer to $\eta^\epsilon(t)$ for $\epsilon > 0$ as a perturbed version of the process $\eta^0(t)$. Let $P^\epsilon(t)$ and $P^0(t)$ denote the transition probability matrices of $\eta^\epsilon(t)$ and $\eta^0(t)$ respectively. Our objective is to analyze the behavior of $\eta^\epsilon(t)$ (or equivalently, that of $P^\epsilon(t)$) as $\epsilon \downarrow 0$ on the time interval $[0, \infty)$.

First, it follows from Proposition 4.2.1 that on any interval of the form $[0, T]$, $\eta^\epsilon(t)$ can be approximated by $\eta^0(t)$. Precisely,

$$\lim_{\epsilon \downarrow 0} \sup_{0 \leq t \leq T} \|P^\epsilon(t) - P^0(t)\| = 0 \quad \forall T < \infty \quad (5.5.2)$$

i.e., the finite dimensional distributions of $\eta^\epsilon(t)$ converge to those of $\eta^0(t)$ uniformly on $[0, T]$. However, as indicated by the example in Section 5.3 and by our results in Chapter IV, the behavior of $\eta^\epsilon(t)$ on the infinite time interval $[0, \infty)$ may differ markedly from that of $\eta^0(t)$.

Following Section 4.2, we shall say that $\eta^\epsilon(t)$ is *regularly perturbed* if

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \|P^\epsilon(t) - P^0(t)\| = 0 \quad (5.5.3)$$

otherwise, we will say that the perturbation is *singular*. In what follows we focus on the singularly perturbed case, since failure of (5.5.3) is symptomatic of the existence of distinct behavior at different time scales.

First, consider the adaptation of the multiple time scale behavior concept to FSMP's.

Definition 5.5.1

We will say that $\eta^\epsilon(t)$ has *well defined behavior at time scale t/ϵ^k* , $k>0$, if there exists a continuous, time-dependent matrix $Y_k(t)$ such that for any $\delta>0$, $T<\infty$,

$$\lim_{\epsilon \downarrow 0} \sup_{\delta < t \leq T} \|P^\epsilon(t/\epsilon^k) - Y_k(t)\| = 0 \quad (5.5.4)$$

Remarks:

1) It is readily verified that the limit matrix $Y_k(t)$ in (5.5.4) must be the transition probability matrix of some FSMP $\eta_k(t)$ taking values in E_0 . Thus (5.5.4) is equivalent to say that $\eta^\epsilon(t/\epsilon^k)$ converges to some FSMP $\eta_k(t)$ as $\epsilon \downarrow 0$ in the sense of finite dimensional distributions.

2) As we will see next, $\lim_{\epsilon \downarrow 0} P^\epsilon(t/\alpha(\epsilon))$ exists for any order function $\alpha(\epsilon)$ ($\alpha: [0, \epsilon_0] \rightarrow \mathbb{R}^+$, $\alpha(0) = 0$ and $\alpha(\cdot)$ continuous and monotone increasing). It turns out, however, that only the limits $Y_k(t)$ for a finite number of positive integers $k=0,1,\dots,m$ are required to construct an asymptotic

approximation to $P^\epsilon(t)$ uniformly valid for $t \geq 0$. We shall call $t, t/\epsilon, \dots, t/\epsilon^m$ the *fundamental or natural time scales* of the process $\eta^\epsilon(t)$.

3) Regularly perturbed processes have trivial time scale behavior. For any order function $\alpha(\epsilon)$

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \|P^\epsilon(t/\alpha(\epsilon)) - \Pi_0^\infty\| = 0 \quad (5.5.5)$$

where Π_0^∞ is the ergodic projection at ∞ of the unperturbed process $\eta^0(t)$.

Proposition 5.5.2

The process $\eta^\epsilon(t)$ is singularly perturbed if and only if the number of ergodic classes at ∞ of the perturbed process $\eta^\epsilon(t)$ is different than that of $\eta^0(t)$ or, equivalently, if $\text{rank } A_0(\epsilon) \neq \text{rank } A_{00}$ for $\epsilon > 0$.

Proof: This is a refrasing of Proposition 4.2.5 in the context of FSMP. The statement in terms of the number of ergodic classes at ∞ follows from the fact that this number equals $\text{nul } A_0(\epsilon)$. □

In Chapter IV we indicated that if a matrix $A_0(\epsilon)$ satisfies the MSST property then $\exp\{A_0(\epsilon)t\}$ has an asymptotic approximation that clearly displays its multiple time scale behavior. We now prove that generators

of FSMP's always satisfy the MSST condition and we construct a uniform asymptotic approximation of singularly perturbed FSMP's based on a hierarchy of aggregated models.

The basic result is the following:

Theorem 5.5.3

Let $\eta^\epsilon(t)$ be a singularly perturbed stochastically continuous FSMP $\eta^\epsilon(t)$ taking values in $E_0 = \{1, 2, \dots, n_0\}$ with transition probability matrix $P^\epsilon(t) = \exp\{A_0(\epsilon)t\}$ and infinitesimal generator $A_0(\epsilon)$ of the form (5.5.1). Denote by $A_k, P_k, k=0, \dots, m$, the sequence of matrices constructed from $A_0(\epsilon)$ as indicated in Section 4.4.1.

Then,

i) A_k and $\Pi_k \triangleq P_0 P_1 \dots P_{k-1}$ are respectively the infinitesimal generator and the ergodic projection at zero of some FSMP $\eta_k(t)$ taking values in E_0 , and

$$\lim_{\epsilon \downarrow 0} \sup_{t \in [\delta, T]} ||P^\epsilon(t/\epsilon^k) - \Pi_k \exp\{A_k t\}|| = 0 \quad (5.5.6)$$

for $\forall \delta > 0, T < \infty$ and $k=1, 2, \dots, m$ (T can be taken equal to ∞ for $k=m$). Furthermore, let $\Pi_k = V_k \cdot U_k$ be the canonical product decomposition of Π_k .

Then,

$$\begin{aligned} \text{ii) } P^\varepsilon(t) &= \exp\{A_0(\varepsilon)t\} = \\ &= \sum_{k=0}^m \exp\{A_k \varepsilon^k t\} - mI + o(1) \end{aligned} \quad (5.5.7)$$

$$= \sum_{k=0}^m \exp\{A_k \varepsilon^k t\} + o(1) \quad (5.5.8)$$

$$= \exp\{A_0 t\} + \sum_{k=1}^m (V_k \exp\{\hat{A}_k \varepsilon^k t\} U_k - \Pi_k) + o(1) \quad (5.5.9)$$

uniformly for $t > 0$, where $\hat{A}_k \triangleq U_k A_k V_k$ is the infinitesimal generator of a stochastically continuous FSMP $\hat{\eta}_k(t)$ taking values in $E_k = \{1, 2, \dots, n_k\}$ and

$$n_k = n_0 - \sum_{p=0}^{k-1} \text{rank } A_p \quad (5.5.10)$$

Proof: The first step is to prove that $A_0(\varepsilon)$ satisfies the MSST property.

We use induction. Suppose that A_0, A_1, \dots, A_ℓ are semistable. Then, by Proposition 4.5.3 the limit

$$\lim_{\varepsilon \downarrow 0} P^\varepsilon(t/\varepsilon^{\ell+1}) = P_0 P_1 \dots P_\ell \exp\{A_{\ell+1} t\} \triangleq P_{\ell+1}(t) \quad \forall t > 0 \quad (5.5.11)$$

is well defined. Clearly, $P_{\ell+1}(t) = I, P_{\ell+1}(t) \geq 0$ and

$P_{\ell+1}(t) \cdot P_{\ell+1}(\tau) = P_{\ell+1}(t+\tau)$ (remember that the projections

$P_j, j=0, 1, \dots, \ell$ commute with each other and with $A_{\ell+1}$). Therefore

$P_{\ell+1}(t)$ is the transition probability matrix of a FSMP and it follows from

Corollary 5.4.8 that $A_{\ell+1}$ must be semistable. Because A_0 is semistable, MSST is proven, and this together with Theorem 4.4.5 gives i). The second part of the theorem follows from (4.4.35) and the fact that, by Theorem 5.4.5

$$\exp\{A_k t\} = I - \Pi_k + V_k \exp\{\hat{A}_k t\} U_k$$

with $A_k = U_k A_k V_k$ being the infinitesimal generator of a stochastically continuous FSMP with $n_k = \text{rank } \Pi_k$ states. Equation (5.5.10) then follows from:

$$\text{rank } \Pi_k = \dim \mathcal{R}(P_0 P_1 \dots P_{k-1}) = n_0 - \sum_{p=0}^{k-1} \text{rank } A_p$$

□

Remarks

1) The matrices $\Pi_k = V_k \cdot U_k$ and A_k satisfy:

$$\text{a) } \Pi_k \Pi_\ell = \Pi_\ell \Pi_k = \Pi_k \quad \underline{k > \ell} \quad (5.5.12)$$

$$\text{b) } U_k \Pi_\ell = U_k \quad \underline{k > \ell} \quad (5.5.13)$$

$$\Pi_\ell V_k = V_k \quad \underline{k > \ell}$$

$$\text{c) } \Pi_k A_\ell = A_\ell \Pi_k = \begin{cases} A_\ell & \underline{k < \ell} \\ 0 & \underline{k > \ell} \end{cases} \quad (5.5.14)$$

2) Part i) implies that the finite dimensional distributions of $\eta^\varepsilon(t/\varepsilon^k)$ converge to those of $\eta_k(t)$ which is, in general, a stochastically discontinuous process.

3) As shown in Section 5.4.2, each of the ergodic projections at zero Π_k , $k=1,2,\dots,m$, determines an aggregation operation performed by collapsing all states that belong to a given ergodic class of Π_k into a single state. If $\eta^\varepsilon(t)$ is aggregated according to the partition specified by Π_k , we get

$$U_k P^\varepsilon(t) V_k = \exp\{\hat{A}_k \varepsilon^k t\} + o(1) \quad (5.5.15)$$

uniformly on $[0, T/\varepsilon^k]$. (This follows using (5.5.13) and (5.4.14) in (5.5.9)). Thus, the aggregation partition specified at stage k isolates transitions between groups of states that are likely to occur over time T/ε^k but not over shorter time intervals. In addition, and to first approximation, these transitions follow a markovian law with rates specified by \hat{A}_k . It is in this sense that we refer to $\hat{\eta}_k(t)$ as an *aggregated model of $\eta^\varepsilon(t)$ valid at time scale t/ε^k* . If such aggregation is not performed, the approximate model for time scale t/ε^k , $\eta_k(t)$, is stochastically discontinuous because transitions that occur at slower time scales look as instantaneous in the limit as $\varepsilon \downarrow 0$.

4) The sequence of aggregated models \hat{A}_k , $k=1,\dots,m$, is a hierarchy. We have already seen that the sequence n_k is non-increasing. Furthermore, (5.5.12) implies that if two rows of Π_λ are equal, the corresponding two

rows in Π_k are also equal $\forall k \geq l$ and therefore if two states are aggregated together at a certain stage then they are also aggregated together at all stages thereafter.

4) In establishing Theorem 5.3.11, the two key properties of FSMP's used are: i) the fact that transition probability matrices are contractions, and ii) positivity. The first property implies MSSNS while positivity assures that the matrices A_k will not have purely imaginary eigenvalues. It is thus clear that any uniformly stable positive system satisfies the MSST condition (see Lemma 3.5.6) and therefore that a result analogous to Theorem 5.5.3 holds for such systems. However, no interpretation of the reduced-order models as some kind of physically interpretable aggregated models is possible in general terms. Any full rank factorization of the projections Π_k (see Lemma 3.2.1) can be used to write a formula analogous (5.5.9) but the interpretation of the reduced-order models \hat{A}_k will depend on the full rank factorization used and on the specific application analyzed.

5) Some aspects of the recursive computation of \hat{A}_k have already been discussed in Section 4.4.5. Thus, for example:

$$\hat{A}_0 = A_0 = A_{00}$$

$$\hat{A}_1 = U_1 A_1 V_1 = U_1 A_{01} V_1$$

$$\hat{A}_2 = U_2 A_2 V_2 = U_2 (A_{02} - A_{01} A_{01}^\#) V_2$$

$$\begin{aligned} \hat{A}_3 = U_3 A_3 V_3 = U_3 & (A_{03} - A_{01} A_{01}^\# A_{02} - A_{02} A_{01}^\# A_{01} + A_{01} A_{01}^\# A_{01}^\# A_{01} \\ & - A_{02} A_{01}^\# A_{02} + A_{02} A_{01}^\# A_{01} A_{01} - A_{01} A_{01}^\# A_{01}^\# A_{02} \\ & - A_{01} A_{01}^\# A_{01}^\# A_{02} - A_{01} A_{01}^\# A_{01}^\# A_{01}^\# A_{01}) V_3 \end{aligned}$$

In addition to the matrix multiplications and additions indicated above, each aggregation stage involves computing a new projection Π_k and a new generalized inverse $A_{k-2}^\#$. Such calculations can be carried out with matrices of increasingly smaller dimension. At the first stage we need to compute Π_1 and $A_0^\#$ and even though these are $(n_0 \times n_0)$ matrices their computation can be decomposed into a set of smaller problems (essentially one per ergodic class of the unperturbed process). At the second stage, Π_2 can be computed as,

$$\Pi_2 = \lim_{t \rightarrow \infty} \Pi_1 \exp\{A_1 t\} = V_1 \left(\lim_{t \rightarrow \infty} \exp\{\hat{A}_1 t\} \right) U_1$$

and therefore only the ergodic projection of the *aggregated* model \hat{A}_1 needs to be calculated. Similarly with the generalized inverse $A_1^\#$,

$$\begin{aligned} A_1^\# &= - \int_0^\infty (e^{A_1 t} - P_1) dt = \int_0^\infty (V_1 e^{\hat{A}_1 t} U_1 + I - \Pi_1 - P_1) dt \\ &= - \int_0^\infty V_1 (e^{\hat{A}_1 t} - \hat{P}_1) U_1 dt = V_1 \hat{A}_1^\# U_1 \end{aligned}$$

which requires only $\hat{A}_1^\# = (\hat{A}_1 + \hat{P}_1)^{-1} - \hat{P}_1$, where \hat{P}_1 is the ergodic projection of the aggregated process \hat{A}_1 . The canonical product decomposition of Π_2 can also be computed from that of Π_1 and that of $\hat{P}_1 = \hat{V}_1 \hat{U}_1$ as follows:

$$\Pi_2 = V_1 \hat{P}_1 U_1 = \underbrace{V_1 \hat{V}_1}_{V_2} \cdot \underbrace{\hat{U}_1 U_1}_{U_2}$$

In summary, the complexity of the computations required decreases with the number of states of the successive aggregated models and they can be implemented in a recursive fashion. We illustrate this procedure with an example in the next section.

5.6 Example

Consider the process $\eta^\varepsilon(t)$ in Figure 5.5. A quick look at the unperturbed version in Figure 5.6 will convince the reader of the singular nature of the perturbation. The ergodic projection at ∞ of $\eta^0(t)$ is given by

$$\Pi_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and its canonical product decomposition is given by:

$$U_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad V_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

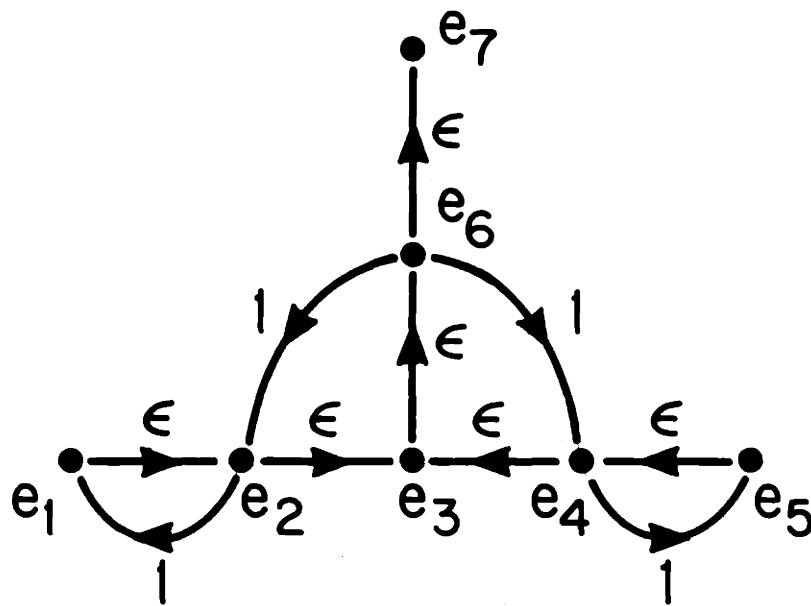


Figure 5.5: The process $\eta^\epsilon(t)$ considered in the example in Section 5.6.

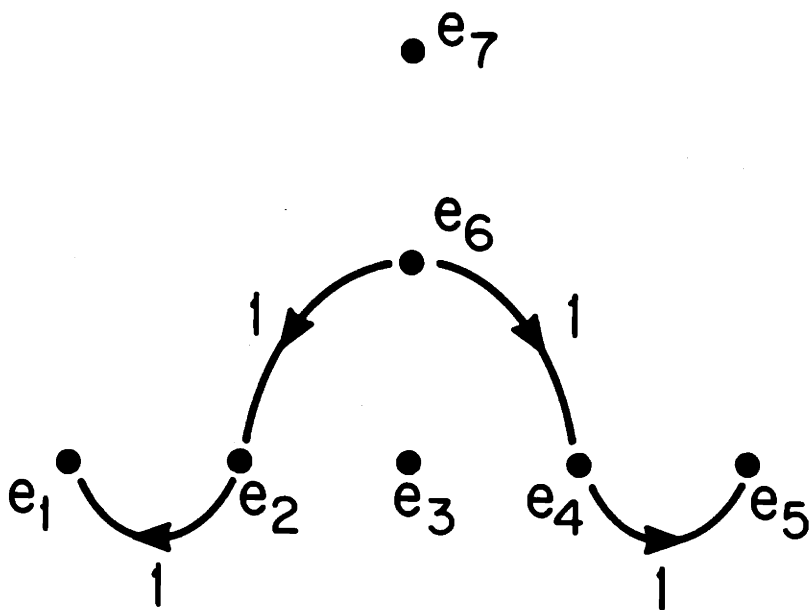


Figure 5.6: The unperturbed process of $\eta^0(t)$.

The ergodic partition at zero has four classes $E_1 = \{1,2\}$, $E_2 = \{3\}$, $E_3 = \{4,5\}$ and $E_4 = \{7\}$ and a transient state $E_T = \{6\}$. The aggregated model $\hat{\eta}_1(t)$ valid at time scale t/ϵ is portrayed in Figure 5.7 and it has the following infinitesimal generator:

$$\hat{A}_1 = U_1 B V_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1/2 & -1 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Notice that the aggregation operation in addition to collapsing $\{1,2\}$ and $\{4,5\}$ into two states also prunes the evanescent state $\{6\}$. To compute next partition, use the ergodic projection of \hat{A}_1 ,

$$\hat{P}_1 = \lim_{t \rightarrow \infty} e^{\hat{A}_1 t} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

to get the canonical product decomposition of $\Pi_2 = V_1 \hat{P}_1 U_1$
 $= (V_1 \hat{V}_1) \cdot (\hat{U}_1 U_1)$ from

$$\hat{U}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

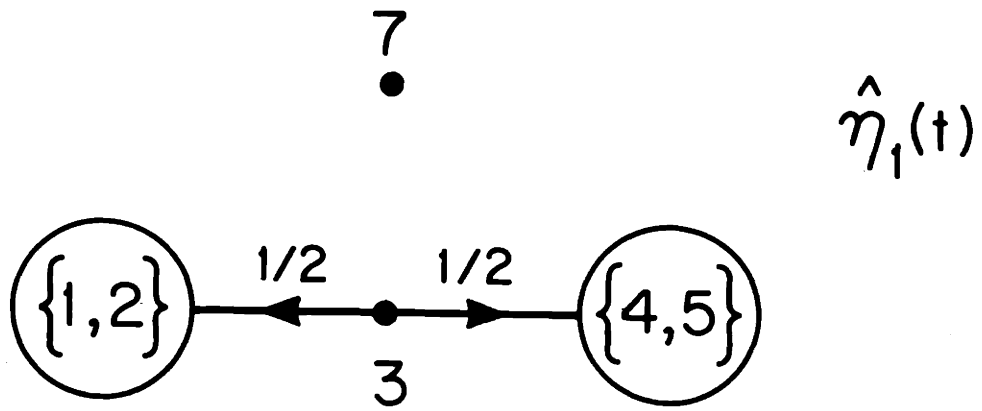


Figure 5.7: Aggregate model valid at time scale t/ϵ .

$$\hat{V}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The resulting partition is:

$$E_1' = \{1,2\}, \quad E_2' = \{4,5\}, \quad E_3' = \{7\}, \quad E_T' = \{3,6\}$$

and the aggregated model valid at time scale t/ϵ^2 , $\hat{\eta}_2(t)$,

represented in Figure 5.8, has generator:

$$\hat{A}_2 = -U_2 BA_0^{\#} BV_2 = \begin{bmatrix} -1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Finally,

$$\hat{P}_2 = \lim_{t \rightarrow \infty} e^{\hat{A}_2 t} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

leads to the next aggregation partition: $E_1'' = \{1,2,3,4,5,6\}$,

$E_2'' = \{7\}$. The aggregated model valid at t/ϵ^3 has rates

$$\hat{A}_3 = -U_3 BA_0^{\#} BV_1 \hat{A}_1^{\#} U_1 BA_0^{\#} BV_3 = \begin{bmatrix} -1/2 & 1/2 \\ 0 & 0 \end{bmatrix}$$

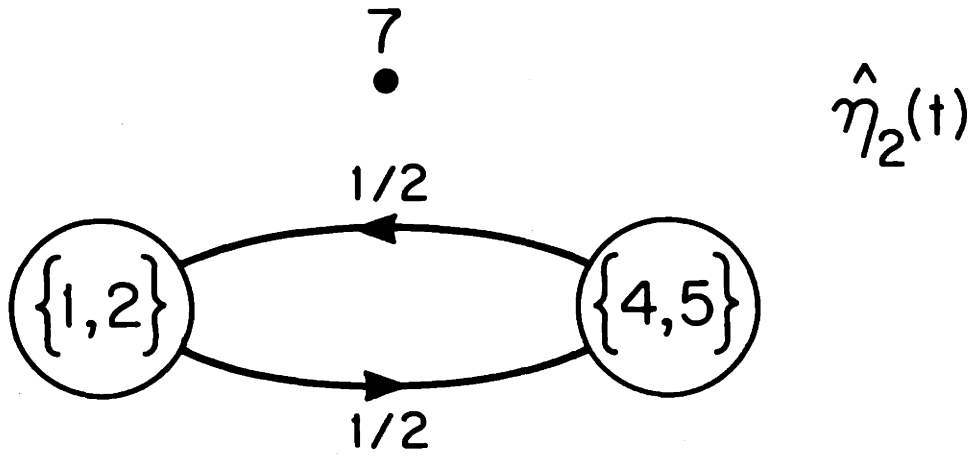


Figure 5.8: Aggregate model valid at time scale t/ϵ^2 .

and it is portrayed in Fig. 5.9. The hierarchy of models ends here because

$$\text{rank } A_0 + \text{rank } \hat{A}_1 + \text{rank } \hat{A}_2 + \text{rank } \hat{A}_3 = 6 = \text{rank } A_0(\varepsilon), \varepsilon > 0$$

This example illustrates how a comparatively complex singularly perturbed FSMP can be asymptotically approximated by a collection of very simple FSMP.

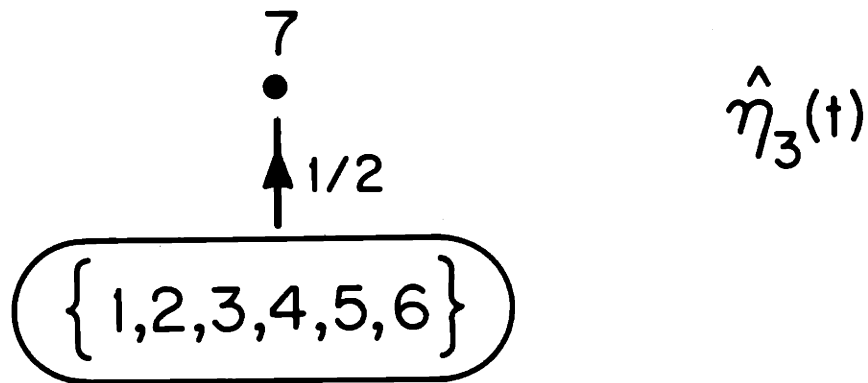


Figure 5.9: Aggregate model valid at time scale t/ϵ^3 .

CHAPTER VI: TIME SCALES AND COHERENCE AREAS
IN POWER SYSTEMS

6.1 Introduction

In the multiple time scale techniques developed in Chapter IV modes are considered fast or slow according to the rate at which they decay to zero which, of course, is specified by the real parts of the eigenvalues. This analysis can adequately be characterized as a decomposition "along the real or σ -axis". There is, however, another possible view of separation of time scales that concentrates on the imaginary parts of the eigenvalues. From this point of view, modes are considered to be fast or slow depending on their oscillation frequency. We will refer to the later as separation of time scales "along the $j\omega$ -axis".

The results presented in Chapter V can be interpreted as saying that for positive systems, time scale separation always takes place along the σ -axis. At least one class of important models, however, exhibit separation of time scales along the $j\omega$ -axis: linearized "swing" equations for electric power networks. A quick look at Figure 6.1, portraying the pole structure of a 16 machine model [Avr 80], will convince the reader that this is indeed the case.

Practitioners in the electric power industry, have for a long time observed that in post-fault transients certain groups of generators "swing together" with an "in-phase" motion of slow frequency [Pod 78].

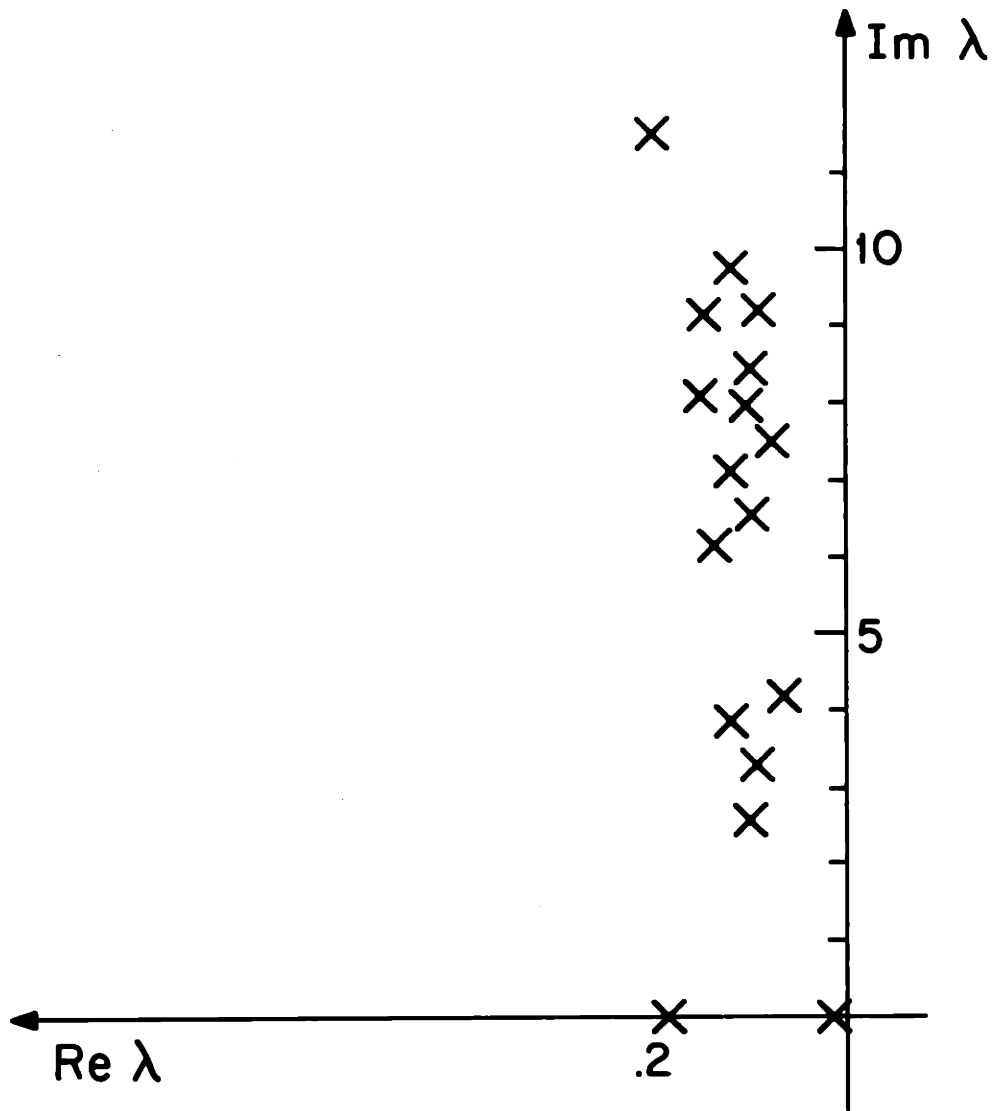


Figure 6.1: Pole structure of a 16 machine linearized power system model.

Each of these groups of generators is considered to be a coherent area and is replaced by a single "equivalent" machine in reduced-order models used in the analysis of post-fault transients [Avr 80]. In practical terms what is required is an algorithm to identify the groups of generators that will swing coherently for a given disruption in the system (i.e. for a given linearized model of the system), and the parameters of the equivalent machines that represent each coherent area.

Implicit in this heuristic reasoning is the notion of a spatial aggregation associated with a time scale separation. Conceptually, this notion is close to the aggregation idea rigorously formalized in Chapter V for FSMP's. In fact, both problems have much in common as we indicate in the sections that follow.

The ideas and results presented in this chapter can also be seen as an extension of the results obtained in Chapter IV to certain types of systems that violate the MSSNS condition. They are important both, for the insight they provide into the coherence area problem, and also because they suggest a way to carry out a more general extension of the asymptotic results developed for the MSST case.

6.2 The Linearized Swing Equations

We review here the model used in the analysis of transients in an n-machine power system. We closely follow [Avr 80].

The intermachine phase variations in a power network are largely determined by the natural frequencies and the mode shapes of the linearized electromechanical model around the stable equilibrium values of the rotor angles δ_i^* and machine speeds ω_i^* . The linearized model is:

$$\Delta \dot{\delta}_i = \Delta \omega_i \quad (6.2.1)$$

$$2H_i \Delta \dot{\omega}_i = -d_i \Delta \omega_i - \sum_{j=1}^n \gamma_{ij} \Delta \delta_j \quad i=1,2,\dots,n \quad (6.2.2)$$

where

$\Delta \delta_i$ = deviation from the equilibrium rotor angle of machine i.

$\Delta \omega_i$ = deviation from the equilibrium machine speed of machine i.

H_i = inertia constant of machine i.

d_i = damping constant of machine i.

γ_{ij} = matrix of interactions between different machines.

The eigenvalues of the above system are of the following three types (see Fig. 6.1):

- i) a zero eigenvalue corresponding to the fact that one of the machines angle can be chosen arbitrary.
- ii) a small negative real eigenvalue corresponding to the collective loss of speed by all the machines, and
- iii) (n-1) pairs of lightly damped oscillatory modes.

Models involving more details would still contain this basic eigenvalue structure with slight modifications mostly in the damping and not in the frequencies of the modes. Furthermore, since it is known that the small damping constants d_i do not significantly alter the frequencies of the oscillatory modes we neglect the damping altogether arriving at the following model:

$$\begin{bmatrix} \dot{\Delta\omega} \\ \dot{\Delta\delta} \end{bmatrix} = \begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix} \begin{bmatrix} \Delta\omega \\ \Delta\delta \end{bmatrix} \quad (6.2.3)$$

where

$$A \triangleq \frac{1}{2} H^{-1} \Gamma \quad (6.2.4)$$

$$H = \text{diag}\{H_1, H_2, \dots, H_n\} \quad (6.2.5)$$

$$\Gamma = (\gamma_{ij}) = V_i V_j B_{ij} \cos(\delta_i^* - \delta_j^*) \quad (6.2.6)$$

It is clear from (6.2.3) that the properties of this model depend exclusively on the structure of the matrix A whose entries are given by:

$$a_{ij} = \frac{1}{2} H_i^{-1} V_i V_j B_{ij} \cos(\delta_i^* - \delta_j^*) \quad i \neq j \quad (6.2.7)$$

$$a_{ii} = - \sum_{j \neq i} a_{ij} \quad (6.2.8)$$

where V_i and V_j are the node voltages, B_{ij} is the imaginary part of the admittance between node i and node j , and δ_i^* , δ_j^* are the stable equilibrium rotor angles of machines i and j . If the network is assumed to be lossless, with no phase-shifters and with the machine angle difference lying in a $\pi/2$ -polytope (i.e., $|\delta_i^* - \delta_j^*| \leq \pi/2$, $\forall i, j$) then A is diagonalizable and all its nonzero eigenvalues σ_i are real and negative. The eigenvalues of (6.2.3) are then $\pm \sqrt{\sigma_i}$ and they are on the imaginary axis close to the slightly damped eigenvalues of (6.2.1) and (6.2.2). The double eigenvalue at zero corresponds now to the fact that the rotor angle and absolute speed of one of the machines are taken as reference.

In addition, and under the conditions stated, all off diagonal terms in A are positive and all its rows sum up to zero. That is, A is also the matrix of transition rates of some FSMP. It is precisely this structure exhibited by the matrix A that will allow us to use the ideas, techniques and results presented in Chapter V in the analysis of coherence phenomena in power systems.

To introduce the notion of time scales in (6.2.3) we will assume that $A=A(\epsilon)$ where ϵ is a small parameter. This small parameter models the fact that that interactions between certain generators may be much weaker than others a feature always present in systems with coherence phenomena. In the sections that follow we will argue that a coherence analysis of the equation:

$$\dot{x}(t) = \begin{bmatrix} 0 & A(\epsilon) \\ I & 0 \end{bmatrix} x(t)$$

is in fact equivalent to the performing the hierarchical aggregation described in Chapter V on $A(\epsilon)$.

For latter reference it is convenient to relate the structure of the eigenprojections and eigennilpotents of the matrix

$$\mathcal{A} = \begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix} \quad (6.2.9)$$

to the corresponding quantities for A .

Proposition 6.2.1

Suppose that A is the diagonalizable matrix of interconnections of a power system introduced above.

Denote by $\lambda_s \neq 0, s=1, \dots, r$ its non-zero eigenvalues and by P_{λ_s} the corresponding eigenprojections. Then $\pm\sqrt{\lambda_s}$ are semisimple eigenvalues values of \mathcal{A} with eigenprojections given by:

$$\mathcal{P}_{\lambda_s}^+ = \frac{1}{2} \begin{bmatrix} P_{\lambda_s} & \sqrt{\lambda_s} P_{\lambda_s} \\ P_{\lambda_s} / \sqrt{\lambda_s} & P_{\lambda_s} \end{bmatrix} \quad (6.2.10)$$

$$\mathcal{P}_{\lambda_s}^- = \frac{1}{2} \begin{bmatrix} P_{\lambda_s} & -\sqrt{\lambda_s} P_{\lambda_s} \\ -P_{\lambda_s} / \sqrt{\lambda_s} & P_{\lambda_s} \end{bmatrix} \quad (6.2.11)$$

Let the zero eigenvalue of A , $\lambda_0=0$, have multiplicity m_0 , then zero is a non-semisimple eigenvalue of \mathcal{A} with algebraic multiplicity equal to $2m_0$ and the corresponding eigenprojection and eigennilpotent are given by:

$$\mathcal{P}_{\lambda_0} = \begin{bmatrix} P_{\lambda_0} & 0 \\ 0 & P_{\lambda_0} \end{bmatrix} \quad (6.2.12)$$

$$\mathcal{D}_0 = \begin{bmatrix} 0 & 0 \\ P_{\lambda_0} & 0 \end{bmatrix} \quad (6.2.13)$$

Proof: Note first that by the spectral decomposition properties of the matrices P_{λ_s} , $s=0,1,\dots,r$, we have:

$$P_{\lambda_0} + \sum_{s=1}^r (P_{\lambda_s}^+ + P_{\lambda_s}^-) = I \quad (6.2.14)$$

and

$$P_{\lambda_i} P_{\lambda_j} = \delta_{ij} P_{\lambda_i} \quad i \neq j \quad (6.2.15)$$

Also,

$$\mathcal{A} P_{\lambda_s}^{\pm} = \pm \sqrt{\lambda_s} P_{\lambda_s} \quad (6.2.16)$$

and

$$\mathcal{A}\mathcal{P}_{\lambda_0} = \mathcal{D}_{\lambda_0} \tag{6.2.17}$$

Thus

$$\mathcal{A} = \mathcal{D}_{\lambda_0} + \sum_{s=1}^r \left(\sqrt{\lambda_s} \mathcal{P}_{\lambda_s}^+ - \sqrt{\lambda_s} \mathcal{P}_{\lambda_s}^- \right) \tag{6.2.18}$$

and because \mathcal{D}_{λ_0} commutes with $\mathcal{P}_{\lambda_s}^{\pm}$ the uniqueness property of the spectral decomposition (6.2.18) gives the desired result. \square

Remark: The above proposition indicates that it is not possible to conduct a multiple time scale analysis of

$$\mathcal{A}(\varepsilon) = \begin{bmatrix} 0 & A(\varepsilon) \\ I & 0 \end{bmatrix}$$

along the lines followed in Chapter IV because neither $\mathcal{A}(0)$ nor $\mathcal{A}(\varepsilon)$ have semisimple nullstructure. Furthermore, we will see that even if the fixed double pole at zero is excluded the remaining matrix does not satisfy the MSST condition. However, given the structure of $\mathcal{A}(\varepsilon)$ it is nevertheless possible to develop some analogous results for this special case of a non MSSNS system and, in addition, to relate these results to the coherence area analysis that is usually performed in practical situations.

In the next section we carry out a complete analysis of a simple example and in Section 6.4 we outline the approach for the general case.

6.3 A Motivating Example

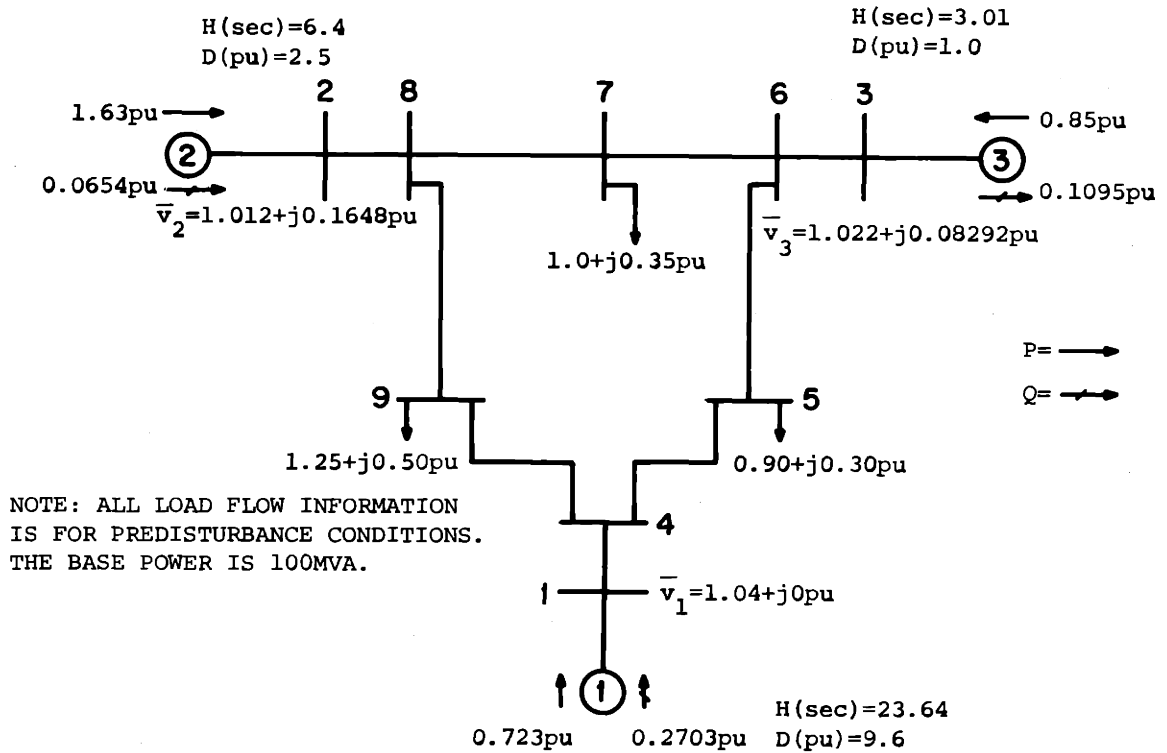
Consider the three-machine example shown in Figure 6.2. For the given numerical values the undamped linearized model turns out to be:

$$A = \begin{bmatrix} -14.3 & 5.5 & 8.8 \\ 7.1 & -42.2 & 35.1 \\ 5.8 & 81.5 & -87.3 \end{bmatrix} \quad (6.3.1)$$

The reader will notice that A is the matrix of transition rates of a FSMP with states 2 and 3 strongly interacting and with weaker connections to state 1. Based on this observation we can attempt to decompose A into a form $A_0 + \epsilon B$ which separates the strong and weak interactions. One such decomposition is:

$$A(\epsilon) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -35.1 & 35.1 \\ 0 & 81.5 & -81.5 \end{bmatrix} + \epsilon \begin{bmatrix} -71.5 & 27.5 & 44 \\ 35.5 & -35.5 & 0 \\ 29 & 0 & -29 \end{bmatrix}$$

$$\stackrel{\Delta}{=} A_0 + \epsilon B \quad (6.3.2)$$



Line#	From	To	R(pu)	x(pu)	B/2(pu)
1	1	4	0	0.0567	0
2	4	5	0.017	0.0567	0.079
3	5	6	0.039	0.170	0.179
4	3	6	0	1.0586	0
5	6	7	0.0119	0.1008	0.1045
6	7	8	0.0085	0.072	0.0745
7	8	2	0	0.0625	0
8	8	9	0.032	0.161	0.153
9	9	4	0.01	0.085	0.088

Figure 6.2: A three machine example. (Taken from [Avr 80]).

with $\epsilon = 0.2^*$

The eigenvalues of $A_0(\epsilon)$ are $\lambda_0 = 0$, $\lambda_1(\epsilon) = -116.6 + o(1)$ and $\lambda_2(\epsilon) = -104 \cdot \epsilon + o(\epsilon)$ and those of

$$\mathcal{A}_0(\epsilon) = \begin{bmatrix} 0 & A_0(\epsilon) \\ I & 0 \end{bmatrix} \quad (6.3.3)$$

are $\lambda_0 = 0$, $\lambda_{1,2}(\epsilon) = \pm 10.8i + o(1)$ and $\lambda_{3,4}(\epsilon) = \pm 10.2 \sqrt{\epsilon} i + o(\sqrt{\epsilon})$. The coherence area problem, as usually stated, is to construct a reduced-order model of (6.3.3) by combining different generators on a coherence area in such a way that the simplified model approximately retains the slow-eigenvalues eigenstructure, i.e., the eigenvalues $\lambda_{3,4}(\epsilon) = \pm 10.2 \sqrt{\epsilon} + o(\sqrt{\epsilon})$.

Consider now aggregating $A_0(\epsilon)$ using the methodology developed in Chapter V. The aggregated model valid at time scale t/ϵ is found to be:

$$\hat{A}_1 = U_1 B V_1 = \begin{bmatrix} -71.5 & 71.5 \\ 33.5 & -33.5 \end{bmatrix} \quad (.3.4)$$

* The problem of how to decompose a given matrix into its weak and strong interactions is not treated in this thesis. The reader is referred to Chapter VIII for our point of view on this important problem and some suggestions for future work in this area. For the case in hand, let us just say that the basic point to be made with the example does not depend on the detailed values assigned to the entries of A_0 and B, but only on separating states 2 and 3 from state 1 in the unperturbed matrix A_0 .

and it has two eigenvalues $\hat{\lambda}_0=0$ and $\hat{\lambda}_1=-104$. Notice that

$$\hat{\mathcal{A}}_1(\epsilon) = \begin{bmatrix} 0 & \epsilon \hat{A}_1 \\ I & 0 \end{bmatrix} \quad (6.3.5)$$

are the swing equations of an electric power network with two generators which, except for terms of order $o(\sqrt{\epsilon})$, has as eigenvalues of $\mathcal{A}_1(\epsilon)$. Furthermore, as indicated in Section 5.5 the order reduction from $\mathcal{A}_0(\epsilon)$ to $\mathcal{A}_1(\epsilon)$ is readily interpreted as consolidating generators 2 and 3 of $\mathcal{A}_0(\epsilon)$ into generator 2 of $\mathcal{A}_1(\epsilon)$. Let us emphasize that it is the structure of the matrix of interconnections $A(\epsilon)$ what permits us to *directly* use the results on FSMP aggregation in the coherence area problem.

This reasoning thus "solves" the coherence problem as usually stated: Given the linearized swing equation of a power network find a reduced-order model which preserves the slow eigenstructure by consolidating groups of generators into an equivalent generator. It also shows that in general it is possible to think of a hierarchy of coherence area models corresponding to the hierarchical aggregation of $A_0(\epsilon)$.

The link between $\mathcal{A}_0(\epsilon)$ and the coherence area model $\hat{\mathcal{A}}_1(\epsilon)$ as establish above is nevertheless incomplete because the question as to what is the relationship between $\exp\{\mathcal{A}_0(\epsilon)t\}$ and $\exp\{\hat{\mathcal{A}}_1(\epsilon)t\}$

is not addressed. Based on the aggregation results presented in Section 5.5 we could expect that there exists an aggregation operation (i.e., a pair of matrices U_1 and V_1) such that

$$U_1 \exp\{A_0(\epsilon)t\} V_1 \approx \exp\{\hat{A}_1(\epsilon)t\} \quad (6.3.6)$$

with the precise nature of this approximation to be determined.

Notice that (6.3.6) cannot be an asymptotic approximation of the type developed in Chapter IV because both $A_0(\epsilon)$ and $A_1(\epsilon)$ fail to satisfy the MSSNS condition and therefore both sides of (6.3.6) grow unbounded in norm as $\epsilon \downarrow 0$. Instead, as we will now see, each entry of $\exp\{A_1(\epsilon)t\}$ is found to be the dominant term of a series (in general a Laurent series) expansion of the corresponding entry in $U_1 \exp\{A_0(\epsilon)t\} V_1$.

First we need to establish some notation. Let

$$\begin{aligned} \lambda_0 &= 0 \\ \lambda_1(\epsilon) &= \epsilon\lambda_1 + o(\epsilon) \\ \lambda_2(\epsilon) &= \lambda_2 + o(1) \end{aligned}$$

be the three eigenvalues of $A_0(\epsilon)$ and denote by $P_{\lambda_0}(\epsilon)$, $P_{\lambda_1}(\epsilon)$ and $P_{\lambda_2}(\epsilon)$ the corresponding eigenprojections.

It follows from Section 3.4 that

$$P_{\lambda_2}(\epsilon) = P_{\lambda_2} + \sum_{k=1}^{\infty} \epsilon^k P_{\lambda_2}^{(k)}$$

$$P_0(\epsilon) \triangleq P_{\lambda_0}(\epsilon) + P_{\lambda_1}(\epsilon) = P_0 + \sum_{k=1}^{\infty} \epsilon^k P_0^{(k)}$$

where P_{λ_2} and P_0 are respectively the eigenprojections for the eigenvalues $\lambda_0=0$ and λ_2 of $A_0(0) = A_0$. Furthermore, it follows from [Kat 60, p. 69-70] that

$$\lim_{\epsilon \downarrow 0} P_{\lambda_0}(\epsilon) = P_{\lambda_0}$$

is well defined and therefore we have

$$P_0 = P_{\lambda_0} + P_{\lambda_1} \tag{6.3.7}$$

We are now ready to write the spectral decomposition of $\exp\{A_0(\epsilon)t\}$:

$$\exp\{A_0(\epsilon)t\} = \mathcal{D}_0(\epsilon)t +$$

$$+ \sum_{k=1}^2 \left[\begin{array}{cc} P_{\lambda_k}(\epsilon) \cos \sqrt{\lambda_k(\epsilon)} t & \sqrt{\lambda_k(\epsilon)} P_{\lambda_k}(\epsilon) \cos \sqrt{\lambda_k(\epsilon)} t \\ \frac{P_{\lambda_k}(\epsilon)}{\sqrt{\lambda_k(\epsilon)}} \sin \sqrt{\lambda_k(\epsilon)} t & P_{\lambda_k}(\epsilon) \cos \sqrt{\lambda_k(\epsilon)} t \end{array} \right]$$

$$\tag{6.3.8}$$

where

$$\mathcal{A}_0(\varepsilon) = \begin{bmatrix} 0 & 0 \\ P_{\lambda_0}(\varepsilon) & 0 \end{bmatrix}$$

is the eigennilpotent for the zero eigenvalue of $\mathcal{A}_0(\varepsilon)$. Retaining the dominant term in ε , in each entry of the above matrices, we have:

$$\begin{aligned} \exp\{\mathcal{A}_0(\varepsilon)t\} \approx & \begin{bmatrix} 0 & 0 \\ P_{\lambda_0} & 0 \end{bmatrix} t + \\ & + \begin{bmatrix} P_{\lambda_1} \cos \sqrt{\varepsilon\lambda_1} t & \sqrt{\varepsilon\lambda_1} P_{\lambda_1} \cos \sqrt{\varepsilon\lambda_1} t \\ \frac{P_{\lambda_1}}{\sqrt{\varepsilon\lambda_1}} \sin \sqrt{\varepsilon\lambda_1} t & P_{\lambda_1} \cos \sqrt{\varepsilon\lambda_1} t \end{bmatrix} + \\ & + \begin{bmatrix} P_{\lambda_2} \cos \sqrt{\lambda_2} t & \sqrt{\lambda_2} P_{\lambda_2} \cos \sqrt{\lambda_2} t \\ \frac{P_{\lambda_2}}{\sqrt{\lambda_2}} \sin \sqrt{\lambda_2} t & P_{\lambda_2} \cos \sqrt{\lambda_2} t \end{bmatrix} \end{aligned} \tag{6.3.9}$$

Notice that this approximation still retains the complete eigenstructure of $\mathcal{A}_0(\varepsilon)$: a double pole at zero a pair of fast eigenvalues and a pair of slow eigenvalues. The approximation consists

in keeping only the dominant terms in ϵ in the eigenvalues $\lambda_s(\epsilon)$ and in the eigenprojections $P_{\lambda_s}(\epsilon)$.

To isolate the slow modes in (6.3.9) we use the aggregation operation determined by the eigenprojection for the zero eigenvalue of $\mathcal{A}_0(0)$, i.e.,

$$P_0 = \begin{bmatrix} P_0 & 0 \\ 0 & P_0 \end{bmatrix} = \begin{bmatrix} V_1 & 0 \\ 0 & V_1 \end{bmatrix} \cdot \begin{bmatrix} U_1 & 0 \\ 0 & U_1 \end{bmatrix} \triangleq U_1 \cdot V_1 \quad (6.3.10)$$

where $P_0 = V_1 \cdot U_1$ is the canonical product decomposition introduced in Section 5.4. We thus have:

$$U_1 \exp\{\mathcal{A}_0(\epsilon)t\} V_1 \approx \begin{bmatrix} 0 & 0 \\ U_1 P_{\lambda_0} V_1 & 0 \end{bmatrix} t + \begin{bmatrix} U_1 P_{\lambda_1} V_1 \cos\sqrt{\epsilon\lambda_1} t & \sqrt{\epsilon\lambda_1} U_1 P_{\lambda_1} V_1 \cos\sqrt{\epsilon\lambda_1} t \\ \frac{U_1 P_{\lambda_1} V_1}{\sqrt{\epsilon\lambda_1}} \sin\sqrt{\epsilon\lambda_1} t & U_1 P_{\lambda_1} V_1 \cos\sqrt{\epsilon\lambda_1} t \end{bmatrix} \quad (6.3.11)$$

To establish that the two matrices in the right-hand side of

(6.3.11) equal $\exp\{\hat{\mathcal{A}}_1(\epsilon)t\}$ we only need to show that zero and λ_1

are eigenvalues of \hat{A}_1 and that $U_{1\lambda_0}^P V_1$ and $U_{1\lambda_1}^P V_1$ are their respective eigenprojections. First,

$$\begin{aligned} (U_{1\lambda_i}^P V_1) \cdot (U_{1\lambda_j}^P V_1) &= U_{1\lambda_i}^P (P_{\lambda_0} + P_{\lambda_1}) P_{\lambda_j} V_1 = \\ &= \delta_{ij} \cdot U_{1\lambda_i}^P V_1 \end{aligned} \tag{6.3.12}$$

and

$$U_{1\lambda_0}^P V_1 + U_{1\lambda_1}^P V_1 = U_{10}^P V_1 = I$$

Thus, $U_{1\lambda_0}^P V_1$ and $U_{1\lambda_1}^P V_1$ are mutually annihilating projections

that sum up to the identity. Further,

$$\begin{aligned} \hat{A}_1 &= \lim_{\epsilon \rightarrow 0} U_1 \cdot \frac{P_0(\epsilon) A_0(\epsilon)}{\epsilon} V_1 \\ &= U_1 \cdot \lim_{\epsilon \rightarrow 0} \left[\frac{1}{\epsilon} (P_{\lambda_0}(\epsilon) + P_{\lambda_1}(\epsilon)) A_0(\epsilon) \right] \cdot V_1 \\ &= U_1 \cdot \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (\epsilon \lambda_1(\epsilon) P_{\lambda_1}(\epsilon)) \cdot V_1 \\ &= \lambda_1 \cdot U_{1\lambda_1}^P V_1 \end{aligned} \tag{6.3.13}$$

Equation (6.3.13) is the spectral decomposition of \hat{A}_1 and we finally conclude

$$U_1 \exp\{\mathcal{A}_0(\epsilon)t\} V_1 \approx \exp\{\hat{\mathcal{A}}_1(\epsilon)t\} \tag{6.3.14}$$

in the sense that each entry of the aggregated model in the right hand side coincides with the first order term in ϵ of the corresponding entry of the left hand side matrix.

Notice that this asymptotic approximation, in contrast with those developed in Chapter IV, does not imply well defined time scale behavior in the sense of Definition 4.2.3. The limit

$$\lim_{\epsilon \downarrow 0} \exp\left\{ \mathcal{D}_0(\epsilon) t / \alpha(\epsilon) \right\} \quad (6.3.15)$$

does not exist for any order function $\alpha(\epsilon) \rightarrow 0$ among other things because of the double pole at zero that produces the nilpotent block

$$\mathcal{D}_0(\epsilon) t = \begin{bmatrix} 0 & 0 \\ P_{\lambda_0}(\epsilon) & 0 \end{bmatrix} t \quad (6.3.16)$$

This mode with linear growth in t is a characteristic feature of the linearized swing equations that both the complete and the aggregated, coherence-based model should (and do) exhibit. Because this unavoidable, common unstable mode impedes uniform asymptotic approximations it is convenient to deal with it separately, and to focus our attention on the oscillatory modes.

It follows from (6.3.8) and (6.3.11) that

$$\begin{aligned}
 \mathcal{U}_1 [\exp\{ \mathcal{A}_0(\epsilon)t \} - \mathcal{D}_0(\epsilon)t] \mathcal{V}_1 \approx \\
 \exp\{ \hat{\mathcal{A}}_1(\epsilon)t \} - \hat{\mathcal{D}}_0 t = \\
 \left[\begin{array}{cc}
 U_{1P} \lambda_1 V_1 \cos \sqrt{\epsilon \lambda_1} t & \sqrt{\epsilon \lambda_1} U_{1P} \lambda_1 V_1 \cos \sqrt{\epsilon \lambda_1} t \\
 \frac{U_{1P} \lambda_1 V_1}{\sqrt{\epsilon \lambda_1}} \sin \sqrt{\epsilon \lambda_1} t & U_{1P} \lambda_1 V_1 \cos \sqrt{\epsilon \lambda_1} t
 \end{array} \right]
 \end{aligned}
 \tag{6.3.17}$$

isolating in this way the slow frequency modes associated with the coherence area approximation. Even with this modification, however, there is no well defined behavior at the time scale at which the slow oscillations take place, i.e., $t/\sqrt{\epsilon}$. This is so because the entry (2,1) in (6.3.17) implies

$$\lim_{\epsilon \downarrow 0} \left| \exp\{ \mathcal{A}_0(\epsilon)t/\sqrt{\epsilon} \} - \mathcal{D}_0(\epsilon)t/\sqrt{\epsilon} \right| = \infty$$

for any $t > 0$, a manifestation of the observed fact [Cha 81] that the slow oscillations due to coherence phenomena have large amplitudes.

This example illustrates that coherence area models are a natural extension of the aggregation results presented in Chapter V and, more importantly, the example gives a clear indication of the nature of the approximation: it consists of retaining the dominant term *in each entry* of the transition matrix. This approximation is

peculiar in two ways. First, it maintains the complete mode structure independently of the order in ϵ of the different modes, and, in this sense, it is a very natural approximation. On the other hand, it is *not* a uniform approximation in t , i.e., no matter how small ϵ is, the error gets arbitrarily large as $t \rightarrow \infty$ in some modes. Thus, care must be exercised in evaluating the interval of validity of the approximation.

6.4 Multiple Time Scale Behavior of Linearized Swing Equation

As suggested by the example above, linearized swing equations are a class of non-MSSNS systems for which the results developed in Chapter IV extend in a natural way. We now discuss this point in more detail.

Suppose that $\mathcal{A}_0(\epsilon)$ is a $(2n \times 2n)$ matrix of the form:

$$\mathcal{A}_0(\epsilon) = \begin{bmatrix} 0 & A_0(\epsilon) \\ I & 0 \end{bmatrix} \quad (6.4.1)$$

with $A_0(\epsilon)$ an $(n \times n)$ matrix that satisfies the MSSNS condition. Let $P_k(\epsilon)$, $Q_k(\epsilon)$ and $A_k(\epsilon)$, $k=0,1,\dots,m$, be the sequence of matrices constructed from $A_0(\epsilon)$ as in Section 4.4.1, i.e.,

$$A_0(\epsilon) = \sum_{k=0}^m \epsilon^k Q_k(\epsilon) A_k(\epsilon) \quad (6.4.2)$$

and $P_k(\epsilon) = I - Q_k(\epsilon)$. The result below generalizes Theorem 4.4.4 to systems of the type (6.4.1)

Theorem 6.4.1

Let $\mathcal{A}_0(\epsilon)$ be as in (6.4.1) with $A_0(\epsilon)$ MSSNS.

Then,

$$\begin{aligned} \exp\{\mathcal{A}_0(\epsilon)t\} &= \sum_{k=0}^m \mathcal{Q}_k(\epsilon) \exp\{\mathcal{A}_k(\epsilon)\epsilon^k t\} + \mathcal{P}_0(\epsilon) \dots \mathcal{P}_m(\epsilon) \\ &= \sum_{k=0}^m \exp\{\mathcal{Q}_k(\epsilon) \mathcal{A}_k(\epsilon) \epsilon^k\} - mI \\ &= \prod_{k=0}^m \exp\{\mathcal{Q}_k(\epsilon) \mathcal{A}_k(\epsilon) \epsilon^k t\} \\ &= \exp\left\{ \sum_{k=0}^m \mathcal{Q}_k(\epsilon) \mathcal{A}_k(\epsilon) \epsilon^k t \right\} \end{aligned} \quad (6.4.3)$$

where,

$$\mathcal{P}_k(\epsilon) = \begin{bmatrix} P_k(\epsilon) & 0 \\ 0 & P_k(\epsilon) \end{bmatrix}$$

$$\mathcal{Q}_k(\epsilon) = I - \mathcal{P}_k(\epsilon) = \begin{bmatrix} Q_k(\epsilon) & 0 \\ 0 & Q_k(\epsilon) \end{bmatrix} \quad (6.4.4)$$

$$\mathcal{A}_k(\varepsilon) = \begin{bmatrix} 0 & A_k(\varepsilon) \\ P_0(\varepsilon) \dots P_{k-1}(\varepsilon) / \varepsilon^k & 0 \end{bmatrix} \quad (6.4.5)$$

Proof: The proof is based on the same manipulations used in proving Theorem 4.4.4. We sketch here the first step.

Define

$$\mathcal{P}_0(\varepsilon) = \begin{bmatrix} P_0(\varepsilon) & 0 \\ 0 & P_0(\varepsilon) \end{bmatrix}; \quad \mathcal{Q}_0(\varepsilon) = I - \mathcal{P}_0(\varepsilon)$$

then

$$\exp\{\mathcal{A}_0(\varepsilon)t\} = \mathcal{P}_0(\varepsilon)\exp\{\mathcal{A}_0(\varepsilon)t\} + \mathcal{Q}_0(\varepsilon)\exp\{\mathcal{A}_0(\varepsilon)t\}$$

can also be written as,

$$\exp\{\mathcal{A}_0(\varepsilon)t\} = \exp\{\mathcal{P}_0(\varepsilon)\mathcal{A}_0(\varepsilon)t\} - \mathcal{Q}_0(\varepsilon) + \mathcal{Q}_0(\varepsilon)\exp\{\mathcal{A}_0(\varepsilon)t\}$$

Next,

$$\begin{aligned} \mathcal{P}_0(\varepsilon)\mathcal{A}_0(\varepsilon) &= \begin{bmatrix} 0 & P_0(\varepsilon)A_0(\varepsilon) \\ P_0(\varepsilon) & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & A_1(\varepsilon) \\ P_0(\varepsilon) & 0 \end{bmatrix} \triangleq \varepsilon A_1(\varepsilon) \end{aligned}$$

The successive steps and the proof that the different expressions given for $\exp\{\mathcal{A}_0(\varepsilon)t\}$ are equivalent follow immediately as in the MSSNS case. □

As discussed at length in the preceding section, even if $A_0(\epsilon)$ is MSST we cannot have a result similar to Theorem 4.4.5 for $\exp\{\mathcal{A}_0(\epsilon)t\}$. The example analyzed in Section 6.3, however, lends strong support to the following conjecture.

Conjecture 6.4.2

Let $\mathcal{A}_0(\epsilon)$ be as in (6.4.1) with $A_0(\epsilon)$ MSST.

Then

$$\begin{aligned} \exp\{\mathcal{A}_0(\epsilon)t\} &\approx \sum_{k=0}^m \mathcal{Q}_k \exp \left\{ \begin{bmatrix} 0 & \epsilon^k A_k \\ P_0 \dots P_{k-1} & 0 \end{bmatrix} t \right\} - \mathcal{P}_0 \dots \mathcal{P}_m \\ &= \sum_{k=0}^m \exp \left\{ \begin{bmatrix} 0 & \epsilon^k A_k \\ P_0 \dots P_{k-1} & 0 \end{bmatrix} t \right\} - mI \\ &= \prod_{k=0}^m \exp \left\{ \begin{bmatrix} 0 & \epsilon^k A_k \\ P_0 \dots P_{k-1} & 0 \end{bmatrix} t \right\} \end{aligned} \tag{6.4.6}$$

Where the approximation is again to be interpreted as consisting of the dominant term in ϵ , entry by entry. □

Remarks

- 1) For systems with two times scales the above conjecture is readily proved following the same steps used in the example of Section 6.3.

For systems with more than two time scales this method of proof leads to quite involved notation and manipulations. If the conjecture is true, it would be preferable to find a different, more elegant approach.

2) The difficulty in proving Conjecture 6.4.2 is due to its entry by entry nature. Different entries are of different order of magnitude and the result cannot be established using any kind of convergence in any matrix norm. Matrices cannot be manipulated as such but instead each entry must be analyzed individually.

3) If the conjecture is correct then it is possible to define a hierarchy of coherent area models, i.e., a hierarchical sequence of aggregations of the type

$$u_k \exp\{A_o(\epsilon)t\} v_k \approx \exp \left\{ \begin{bmatrix} 0 & \epsilon \hat{A}_k^k \\ I & 0 \end{bmatrix} t \right\} \quad (6.4.7)$$

k=1,2,...,m

exactly as in the FSMP case. The usefulness of this hierarchy for power systems analysis seems, however, quite limited given the fact that the linearized model used most likely loses validity over such long time intervals. It is nevertheless conceptually useful in the situation where the number of slow poles in the system does not match the observed number of coherence areas [Avr 80]: Some very slow poles correspond in this case to coherence areas at a higher level of aggregation.

To conclude this chapter let us emphasize what we see as our contributions in the coherence area analysis for power systems:

1) Our main contribution is of a conceptual and interpretative nature: Starting with a system matrix of the form $A_0 + \epsilon B$ we show that, under certain conditions, the system's evolution exhibits coherence area phenomena. We give a methodology for the construction of an aggregated model, and we give an interpretation of this reduced-order model as an asymptotic approximation of a special kind. It is an asymptotic approximation valid on an entry by entry basis, and it is not uniformly valid over $[0, \infty)$. The nature of the approximation and the lack of a result similar to the MSST case is a manifestation of the practical observation that slow oscillations among coherent areas are of large amplitude if left uncontrolled [Chan 81], behaving as unstable modes on the short run. An observation that our analysis confirms.

2) Beyond that, we provide a clear focus of what must be done to determine groups of coherent generators. Namely, that the problem is totally equivalent to the fundamental question in multiple time scale analysis of how to decompose the matrix of a system into $A_0 + \epsilon B$ with ϵ small and B chosen so that it really captures the singular nature of the perturbation that gives rise to time scale behavior.

The theoretical analysis and perspective developed in this chapter cannot be found in the power systems literature where several authors

(see for example [Pod 78], [Avr 80], [Win 80]), have proposed time-scale based "grouping algorithms" which, although they attempt a decomposition of the sort $A_0 + \epsilon B$, they do so indirectly and without focusing on this as the key point. We felt that our viewpoint can be of help in reinterpreting and possibly modifying these algorithms which are of the utmost practical importance. We should also mention that other authors (see [Sas 81] and references therein) take a different view of coherence, more directly related to the kinds of disturbances experienced by the power system, which does not involve notions of time scale separation. Our analysis belongs more appropriately to the group of authors mentioned first.

Finally, aside from the direct contributions to the coherence area problem, the results of this chapter provide a first indication of how should one proceed to generalize the results in Chapter IV to non-MSSNS systems:

- 1) Unbounded amplitudes as $\epsilon \downarrow 0$ must be accommodated in the definition of well defined multiple time scale behavior and in the asymptotic approximations, and

- 2) Because in general the entries of $\exp\{A_0(\epsilon)t\}$ will be of different orders of magnitude as $\epsilon \downarrow 0$, the asymptotic approximations must be interpreted on an entry by entry basis.

CHAPTER VII: AGGREGATE FILTERING FOR SINGULARLY
PERTURBED FSMP'S

7.1 Introduction

In this chapter we explore the possibility of using aggregated models to simplify the structure of filters for singularly perturbed FSMP's. Let $\rho^\epsilon(t)$ be a FSMP's with infinitesimal generator $A(\epsilon)$ and suppose that we observe $\rho^\epsilon(t)$ in white additive noise, i.e., suppose that we have observations:

$$dy(t) = h(\rho^\epsilon(t))dt + B(\epsilon)dw(t) \quad (7.1.1)$$

where $y(t) \in R^n$, $h(\cdot)$ is a vector valued function and $w(t)$ is a vector of independent, standard Wiener processes. The standard filtering problem consists of computing the a posteriori probabilities:

$$\pi_i(t) = \Pr\{\rho^\epsilon(t)=i | y(\tau), 0 \leq \tau \leq t\}$$

from the observations $y(\tau)$, $0 \leq \tau \leq t$.

Given that we know that the process $\rho^\epsilon(t)$ admits a hierarchical description in terms of a set of simplified models, it is reasonable to attempt to find a filter design that reflects this structure. Such a filter would be composed of a hierarchy of filtering algorithms, each layer running at a different time scale, each using a model that describes $\rho^\epsilon(t)$ with a different degree of aggregation, and with some coordination among the different levels of the algorithm.

An example will clarify this idea. Suppose that $\rho^E(t)$ is the process depicted in Figure 7.1(a). It undergoes fast transitions between states 1 and 2 or 3 and 4, while transitions between $X_1 = \{1,2\}$ and $X_2 = \{3,4\}$ are rare. We can think of a filter in which at a higher level only the rare transitions between X_1 and X_2 are to be detected and which uses the model in Figure 7.1(b) for this purpose. Provided that information about changes between X_1 and X_2 is supplied by the measurements (7.1.1) and at a rate that is high compared to the transition rates $\epsilon\lambda'$ and $\epsilon\lambda''$, this filter will be able to determine quite precisely whether $\rho^E(t) \in X_1$ or $\rho^E(t) \in X_2$. This information can then be fed into a low level fast filter that tries to keep track of the fast transitions in $\rho^E(t)$. The fast filter will use a two state model of the system which at time t has parameters λ_1 and λ_2 or μ_1 and μ_2 according to whether the upper filter estimates $\rho^E(t) \in X_1$ or $\rho^E(t) \in X_2$.

It is clear that other factors, in addition to the multiple time scale structure of the process $\rho^E(t)$, need to be considered when trying to decompose a filtering problem in the way we have just described. In our view, these factors are: *the structure of the measurements, and the rate at which information about certain changes in $\rho^E(t)$ is supplied by the observations $y(t)$.* "Information rate" is a term to be made precise that refers to the magnitude of the signal-to-noise ratio for certain transitions relative to the rate at which these transitions occur, while the "structure of measurements" refers to the relationship between

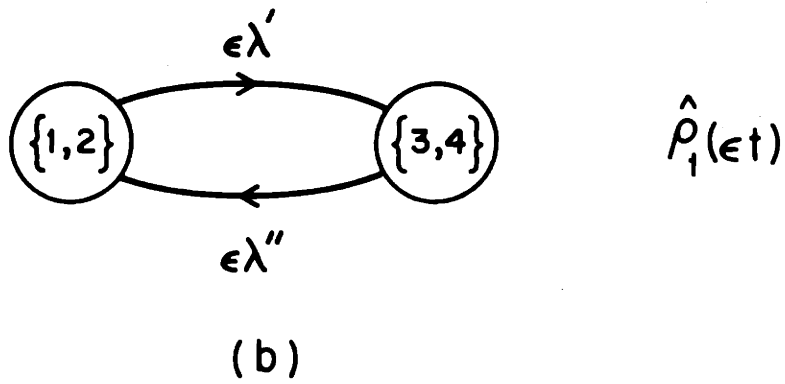
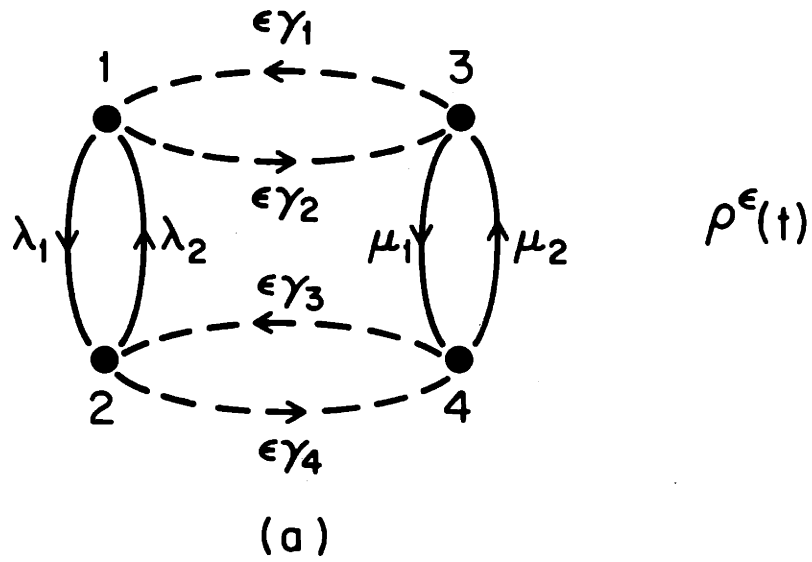


Figure 7.1(a): The process discussed in Section 7.1 and (b) its aggregated model.

the values taken by the observation function $h(p)$ and the partitions specified by the hierarchical structure of $\rho^\varepsilon(t)$.

For example, suppose that we have the following two measurements of the process in Figure 7.1:

$$dy_1(t) = h_1(\rho^\varepsilon(t))dt + b_1dw_1(t) \quad (7.1.2)$$

$$dy_2(t) = h_2(\rho^\varepsilon(t))dt + b_2dw_2(t) \quad (7.1.3)$$

with

$$\begin{aligned} h_1(1) &= h_1(2) \stackrel{\Delta}{=} \alpha_1 \\ h_1(3) &= h_1(4) \stackrel{\Delta}{=} \alpha_2 \end{aligned} \quad (7.1.4)$$

and

$$\begin{aligned} h_2(1) &= h_2(3) \stackrel{\Delta}{=} \beta_1 \\ h_2(2) &= h_2(4) \stackrel{\Delta}{=} \beta_2 \end{aligned} \quad (7.1.5)$$

The structure of the observation function $h_1(\cdot)$ is such that no information at all is supplied by $y_1(t)$ about fast transitions in $\rho^\varepsilon(t)$. On the contrary, $y_2(t)$ mostly carries information about the fast transitions in $\rho^\varepsilon(t)$ and only over longer periods of time it does provide some information about the slow transitions through the different average behavior due to differences in the transition rates λ_1, λ_2 and μ_1, μ_2 . In this case, the structure of the functions $h_1(\cdot)$ and $h_2(\cdot)$ is compatible with the time-scale decomposition of the process $\rho^\varepsilon(t)$ and illustrates the role of the measurement structure in decomposing the filter.

Consider next the notion of information rate, also for this example. The larger the difference is between α_1 and α_2 , the stronger the effect will be of a transition between X_1 and X_2 on the measurement $y_1(t)$. A measure of the signal-to-noise ratio in $y_1(t)$ that we will find convenient is the following

$$K_1 \triangleq \left(\frac{\alpha_1 - \alpha_2}{b_1} \right)^2 \quad (7.1.6)$$

The larger K_1 is the better we can discriminate between $\rho^\varepsilon(t) \in X_1$ and $\rho^\varepsilon(t) \in X_2$. A measure of how well we can track transitions between X_1 and X_2 should involve, in addition, the rate at which such transitions take place. A parameter that will appear in later computations is the ratio between the signal-to-noise ratio K_1 , and the rate of transitions. For the measurement $y_1(t)$ this ratio is

$$\gamma_1 = \frac{K_1}{\varepsilon \lambda} \quad (7.1.7)$$

where λ is some number of the same order of magnitude as λ' and λ'' (for example, $(\lambda' + \lambda'')/2$). We will refer to ratios of this form as information rates. For a given signal-to-noise ratio K_1 , the information rate, γ_1 , supplied by measurement $y_1(t)$ is very high if the transitions to which it refers are rare ($\varepsilon \ll 1$). This high information rate in turn suggests that the higher level estimator for $\rho^\varepsilon(t)$ will perform very well at least for long periods of time in

which no rare transitions take place. Such near perfect performance is a good argument in favor of a hierarchically structured filter. If, on the other hand, the difference $\alpha_1 - \alpha_2$ is very small, say of order ϵ , the information rate γ_1 will be small (also of order ϵ) indicating a poor performance of the higher level filter. In this circumstances a hierarchical filter does not seem advisable.

In general, therefore, a judgement about the hierarchical decomposition of a filter must be based on:

- a) the time scale structure of the process to be estimated;
- b) the structure of the measurements; and
- c) the rate at which information about different transitions is supplied by the measurements.

In addition, a fourth issue must be considered when trying to assess the performance of a hierarchically decomposed filter:

- d) In what sense are we going to judge an estimator to be good or bad?

Hierarchical filters may perform badly at some times. For example when, due to a false alarm at some high level filter, wrong models are used in the lower levels. The frequency of occurrence of such events and their effects have to be quantified, and for that purpose, the standard pointwise minimum mean-square error criteria is not appropriate.

A comprehensive study of the filtering problem for the general case is beyond the scope of this thesis. Instead, our goal in this chapter is to develop some intuition and insight into the issues arising in this problem through the analysis of the example we have just discussed and some related topics motivated by it.

In Section 7.2 we carry out a qualitative analysis of the filtering equations for a two state process with rare transitions. Several additional aspects of this example are developed in Section 7.3 and 7.4. This example motivates the introduction of the signal-to-noise ratio parameter K_1 and of the information rate γ_1 as fundamental quantities, and it illustrates the trade-off involved between detection delays and false alarm probabilities. This trade-off is important in hierarchical filtering because a false alarm (i.e. an erroneous detection) at some high level filter quickly propagates through lower levels and has significant effects on the performance and dynamic behavior of the filter. On the other hand, detection delays at a given level adversely affect overall performance because during this delay, the models used in lower level filters are erroneous.

The qualitative behavior of a hierarchical decomposed filter for the process shown in Fig. 7.1 is discussed in Section 7.5 together with some suggestions for future research on these problems.

7.2 The Optimum Filter Equations: Qualitative Analysis

Let $\rho^\epsilon(t)$ be a singularly perturbed FSMP taking values in $E = \{0, 1, 2, \dots, n\}$ and with infinitesimal generator $A(\epsilon)$. Suppose that the following observation is available.*

$$dy(t) = h(\rho^\epsilon(t))dt + b dw(t) \quad (7.2.1)$$

The vector of a-posteriori probabilities:

$$\pi_i(t) = (\pi(t))_i \triangleq \Pr\{\rho^\epsilon(t)=i | y(\tau), 0 \leq \tau \leq t\} \quad (7.2.2)$$

satisfies the following stochastic differential equation [Lip 78]:

$$d\pi(t) = A(\epsilon)^T \pi(t)dt + \frac{1}{b^2} g(\pi(t)) [dy(t) - \bar{h}(\pi(t))dt] \quad (7.2.3)$$

where:

$$(g(\pi(t)))_i = \pi_i(t) (h(i) - \bar{h}(\pi(t))) \quad (7.2.4)$$

and

$$\bar{h}(\pi(t)) = \sum_{i \in E} h(i) \pi_i(t) \quad (7.2.5)$$

The first term on the right hand side of (7.2.3) corresponds to the evolution of $\pi(t)$ if no measurements were present (the Chapman-Kolmogorov equation) while the second term, the innovations term,

* For simplicity we take the observation to be one-dimensional.

represents the effect on $\pi(t)$ of the measurements. For purposes of analysis, it is convenient to rewrite (7.2.3) as a non-linear diffusion driven by the observation noise. This is easily done by substituting $dy(t)$ in (7.2.3) by its expression (7.2.1) as follows:

$$d\pi(t) = \{A(\epsilon)^T \pi(t) + \frac{1}{2} g(\pi(t)) [h(\rho^\epsilon(t)) - h(\pi(t))]\} dt + \frac{1}{b} g(\pi(t)) dw(t) \quad (7.2.6)$$

Written in this form the filtering equations (7.2.6) are seen to be non-linear diffusion with a drift that depends on the process to be estimated, $\rho(t)$. Furthermore, because $\pi(t)$ is a vector of probabilities, the diffusion (7.2.6) must take place inside the n-dimensional unit simplex, i.e. it is a diffusion in a bounded domain. Notice that in contrast with (7.2.3), equation (7.2.6) is not written in a realizable form (i.e., it cannot be implemented as given) because its coefficients involve the unobservable process $\rho^\epsilon(t)$. This explicit dependence on the process to be estimated, however, makes this equation very useful for purposes of analysis and also gives insight into the filter performance.

To understand better the way filters for FSMP's work we analyze in detail the filtering equations for the two state process indicated in Figure 7.2. For this simple process, the filter is one-dimensional and in its realizable form is given by:

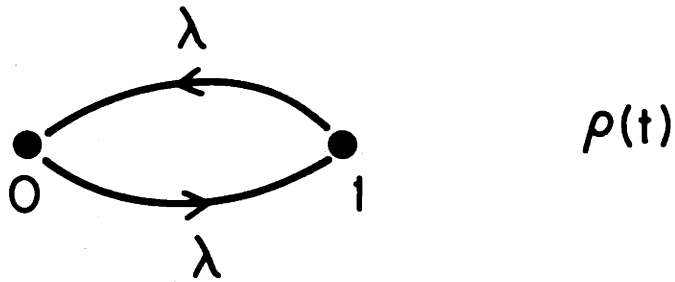


Figure 7.2: The two state process discussed in Section 7.2.

$$d\pi_1(t) = [-\lambda\pi_1(t) + \lambda(1-\pi_1(t))]dt + \frac{1}{b^2} [h(1)-\bar{h}(\pi_1(t))] (dy(t)-\bar{h}(\pi_1(t))dt) \quad (7.2.7)$$

$$\pi_0(t) = 1-\pi_1(t).$$

where

$$\bar{h}(\pi_1(t)) = h(1)\pi_1(t) + h(0)(1-\pi_1(t)) \quad (7.2.8)$$

Written as a diffusion driven by the measurement noise, (7.2.7)

takes the form:

$$d\pi_1(t) = [-\lambda\pi_1(t) + \lambda(1-\pi_1(t))]dt + \frac{h(1)-h(0)}{b^2} \pi_1(t)(1-\pi_1(t)) [h(\rho(t))-\bar{h}(\pi_1(t))]dt + \frac{h(1)-h(0)}{b} \pi_1(t)(1-\pi_1(t))dw(t) \quad (7.2.9)$$

The right hand side of equation (7.2.9) has three terms.

The first term,

$$[-\lambda\pi_1(t)+\lambda(1-\pi_1(t))]dt \quad (7.2.10)$$

corresponds to the a-priori evolution of the probability and it

would be the only one if measurements were not present. The second,

$$\frac{h(1)-h(0)}{b^2} \pi_1(t)(1-\pi_1(t)) [h(\pi(t))-\bar{h}(\pi_1(t))]dt \quad (7.2.11)$$

is a drift term introduced by the measurements. It is the only term affected by a change in the process to be estimated $\rho^\varepsilon(t)$ and therefore it is most important for the filter performance. Finally, the noise term:

$$\frac{h(1)-h(0)}{b} \pi_1(t)(1-\pi(t))dw(t) \quad (7.2.12)$$

reflects the fact that measurement noise affects filter behavior.

An understanding of the qualitative behavior of the filter can be gained from the shape of the drift and noise coefficients. Suppose that $\rho(t)=1$ over the interval of interest. Then $\pi_1(t)$ evolves according to:

$$\begin{aligned} d\pi_1(t) = & [-\lambda\pi_1(t)+\lambda(1-\pi_1(t))]dt \\ & + K^2\pi_1(t)(1-\pi_1(t))^2dt \\ & + K\pi_1(t)(1-\pi_1(t))dw(t) \end{aligned} \quad (7.2.13)$$

where the only parameter appearing in this equation:

$$K^2 \triangleq \left(\frac{h(1)-h(0)}{b} \right)^2 \quad (7.2.14)$$

has a direct interpretation as the signal-to-noise ratio.

Similarly, if $\rho(t)=0$ over a certain interval, then $\pi_1(t)$ evolves according to:

$$\begin{aligned}
 d\pi_1(t) = & [-\lambda\pi_1(t) + \lambda(1-\pi_1(t))]dt \\
 & - K^2\pi_1(t)^2(1-\pi_1(t))dt \\
 & + K\pi_1(t)(1-\pi_1(t))dw(t)
 \end{aligned}
 \tag{7.2.15}$$

In both cases the noise intensity is maximum at $\pi_1(t)=1/2$ (the point of maximum uncertainty where maximum weight is given to the measurements) and it becomes zero at the end points of the interval $[0,1]$ where the drift in both cases points towards the interior of the interval. This is the mechanism that keeps the diffusion confined to the unit interval. In the interior, the drifts of equations (7.2.13) and (7.2.15) have unique stable points at π_1^* and π_0^* respectively that are the solutions of the following equations:

If $\rho(t)=1$

$$\frac{\lambda}{K^2} (1-2\pi_1^*) + \pi_1^*(1-\pi_1^*)^2 = 0
 \tag{7.2.16}$$

If $\rho(t)=0$

$$\frac{\lambda}{K^2} (1-2\pi_0^*) - (\pi_0^*)^2(1-\pi_0^*) = 0
 \tag{7.2.17}$$

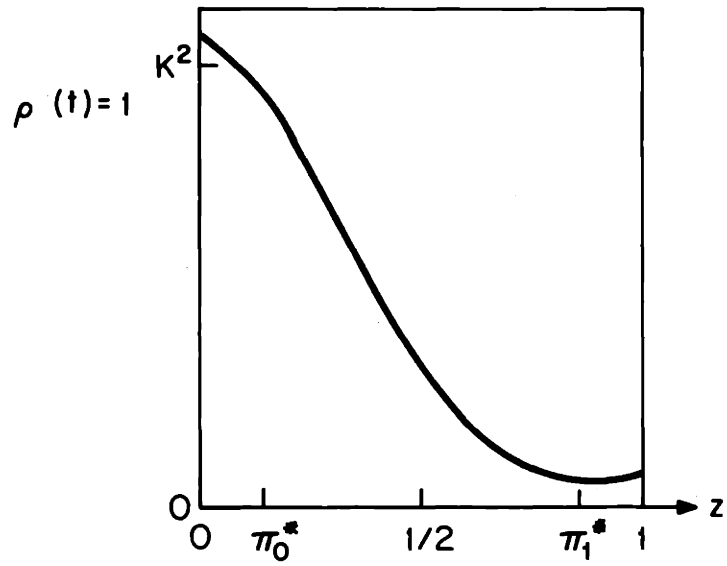
If the parameter

$$\gamma = \frac{K^2}{\lambda}
 \tag{7.2.18}$$

that we term an "information rate", is large, then $\pi_1^* \approx 1$ and $\pi_0^* \approx 0$. If, on the contrary, γ is small $\pi_1^* \approx \pi_0^* \approx 1/2$. Thus, a large γ indicates that the output of the filter will have a tendency to remain close to the correct end point of the interval $[0,1]$ according to the value of $\rho(t)$, while a small value of γ indicates that the measurements do not provide enough information for the filter to be able to track the change in $\rho(t)$, stabilizing instead around the ergodic probabilities of $\rho(t)$. The location of the stable point of a filter diffusion is one of the parameters that we will use in a qualitative analysis of the filter.

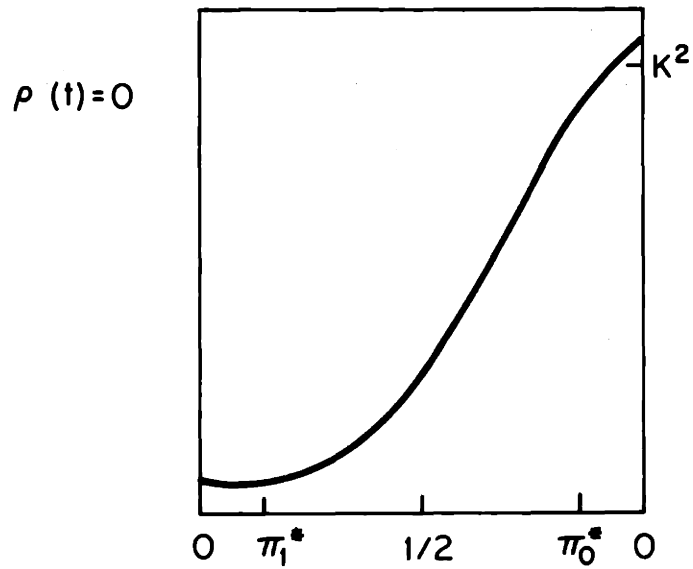
A better idea of the behavior of the filter inside the interval $[0,1]$ and of the effect of a transition in $\rho(t)$ can be gained by plotting the drift coefficients of equations (7.2.13) and (7.2.15). Figure 7.3 portrays the potentials (i.e., the functions whose gradients are the drifts) corresponding to the drift coefficients of the filter for $\rho(t)=0$ and $\rho(t)=1$, and for a value of γ relatively large ($\gamma=10$). Suppose that we have a situation with a quite high signal to noise ratio, $K^2 \gg 1$, and that we are trying to detect transitions that are quite rare, i.e., $\lambda \ll 1$. Suppose further that $\rho(t)$ has stayed at $\rho(t)=1$ for some time and that at $t=0$ there is a transition to $\rho(0^+)=0$. The behavior of the filter will be as follows: Because $K^2 \gg 1$ and the process has remained for some time at $\rho(t)=1$ before switching,

$$\psi(z) = \int_0^z (\lambda(1-2x) + K^2 x(1-x)^2) dx$$



(a) Potential for equation (7.2.13)

$$\psi(z) = \int_0^z (\lambda(1-2x) - K^2 x^2(1-x)) dx$$



(b) Potential for equation (7.2.15)

Figure 7.3: Potential for $\rho(t)=1$ and $\rho(t)=0$.

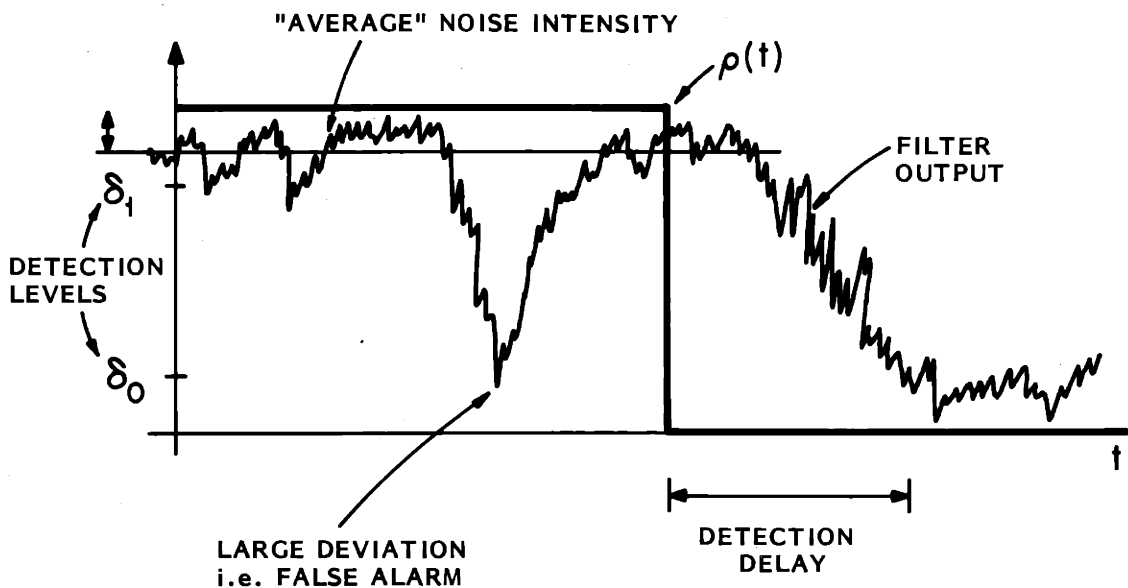
$\pi_1(t)$ will have remained close to $\pi_1^* \approx 1$ on the average. (see Figure 7.3(a)). At $t=0$, when $\rho(t)$ jumps from state zero to state 1, the drift changes from that in Figure 7.3(a) to that in Figure 7.3(b) and, as a consequence, $\pi_1(t)$ will start moving from a neighborhood of π_1^* to a neighborhood of π_0^* , following, on the average, the trajectory determined by the potential well in Figure 7.3(b).

The above discussion together with Figure 7.4(a) suggest the following qualitative measures for the performance of the filter discussed:

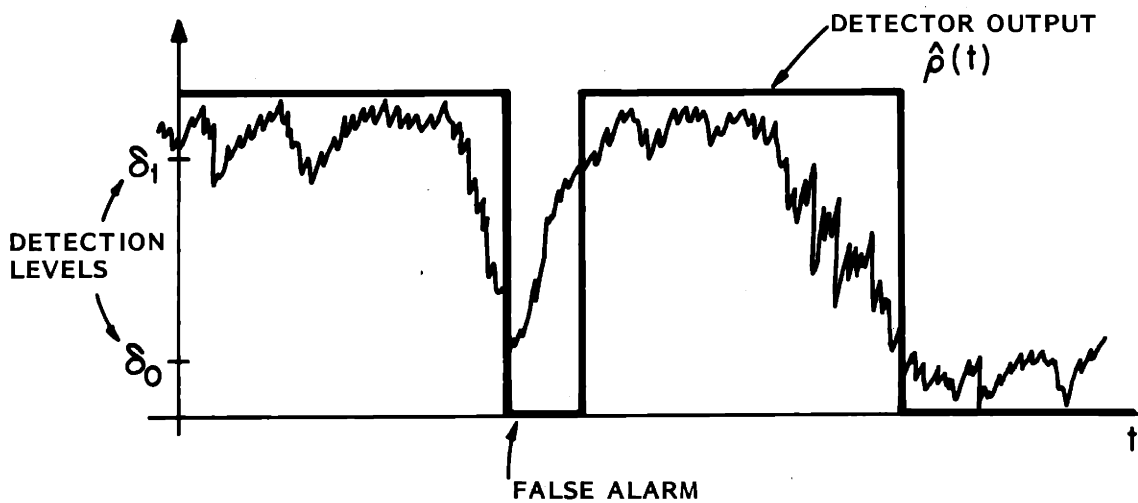
a) FILTER BIAS: The distance between the equilibrium points of the filter diffusion and the extremes of the interval $[0,1]$ or, equivalently, the distance between equilibrium points for different values of $\rho(t)$ give an indication of how well the filter can discriminate between different states of $\rho(t)$ over the long run.

b) NOISE AT EQUILIBRIUM: The variance of the noise around the equilibrium points gives an indication of the noise intensity present at the output of the filter, on the average (i.e., between transitions and far from large deviations).

c) DETECTION DELAY: The time it takes for the filter to evolve from, say, π_0^* to π_1^* following the corresponding transition in $\rho(t)$.



(a) Qualitative measure of filter performance



(b) Detector output for given detection levels δ_0 and δ_1

Figure 7.4: Qualitative performance measures.

d) PROBABILITY OF FALSE ALARM/MEAN TIME BETWEEN FALSE ALARMS:

Given two detection levels, say δ_0 and δ_1 (see Figure 7.4(b)), the probability of a false alarm before a change in $\rho(t)$ occurs and the mean time between such false alarms are well defined quantities. In general, δ_0 and δ_1 are design parameters that are chosen according to a trade-off between delay time to detection and false alarm probabilities.

The filter bias can be obtained from equations (7.2.16) and (7.2.17). If the parameter $\gamma = K^2/\lambda$ is large (high quality measurements) the bias is seen to be:

$$\pi_0^* = 1 - \pi_1^* \approx \lambda/K^2 = 1/\gamma \quad (7.2.19)$$

while for γ small,

$$\pi_0^* = 1 - \pi_1^* \approx 1/2 \quad (7.2.20)$$

In what follows we will mainly study the high quality measurement case (as for example, when trying to detect very rare transitions, $\lambda \ll 1$). The noise intensity at the output of the filter, away from times at which $\rho(t)$ makes transitions or when the filter experiences large excursions, (i.e., false alarms) is approximately given by:

$$\sigma_{av}^2 = (K \cdot \pi_1^* (1 - \pi_1^*))^2 \approx \lambda \quad (7.2.21)$$

As indicated, this average noise intensity is independent of the signal to noise ratio K^2 and it is small only if the transition to be detected are rare.

The calculation of false alarm probabilities, mean time between false alarms and detection delays, and the evaluation of the trade-offs involved in selecting the detection levels δ_0 and δ_1 are more complex and require a detailed analysis that we develop in the sections that follow, starting with the standard hypothesis testing case, i.e., with $\lambda=0$.

7.3 The Hypothesis Testing Problem

In this section we study the detection problem for $\lambda=0$, as a preliminary step in order to understand the trade-off involved in the detection of rare events. For $\lambda=0$, $\rho(t) = \rho(0)$, and suppose, as in Section 7.2, that we observe:

$$dy(t) = h(\rho(0))dt + b dw(t) \quad (7.3.1)$$

Suppose further that we use the filter (7.2.7) with $\lambda=0$ to compute $\pi_1(t) = \Pr\{\rho(0)=1|y(\tau), 0 \leq \tau \leq t\}$ and that we want to decide between $\rho(0)=0$ and $\rho(0)=1$ according to whether the level δ_0 or δ_1 is first crossed (see Fig. 7.5).

The detection scheme just described performs a sequential test of the kind introduced by Wald [Wal 47]. In his work and in subsequent research in hypothesis testing problems with the length of observations not fixed a priori, performance analysis has typically been based on a trade-off between two parameters: *probability of an erroneous decision and mean time to decision*. In our case, this means

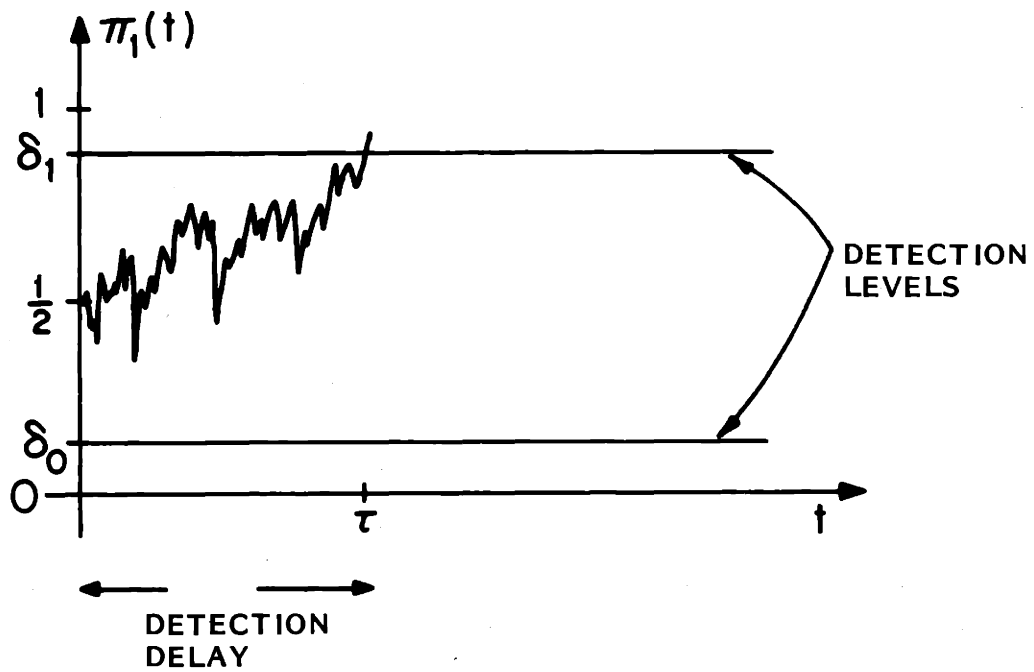


Figure 7.5: The Hypotesis Testing Problem.

that the filter output $\pi_1(t)$ would be observed only until it exits the interval $[\delta_0, \delta_1]$ when a decision would be made. We develop this classical analysis for the proposed filter in Section 7.3.2.

Our objective, however, is not to analyze the hypothesis testing problem *per se* but rather to expose the dependence on several critical parameters and to use this as a first step when we view the problem as the limit of a filtering problem with very rare transitions. From this perspective, it becomes necessary to analyze the evolution of $\pi_1(t)$ over the infinite time interval $[0, \infty)$. Specifically, we are interested in the probability and the frequency of occurrence of *large deviations* in $\pi_1(t)$. For example, assuming that $\rho(0)=1$ and that the level δ_1 has already been reached, what is the probability that $\pi_1(t)$ will cross the level δ_0 ? If this probability is positive, what is the expected time between such large excursions?.

To our knowledge these questions have not been addressed in the sequential analysis literature. We develop them in Section 7.3.3. The results obtained there are useful because, as we will see in Section 7.4, these large excursions produce false alarms when the filter $\pi_1(t)$ is used to detect rare transitions.

Given the symmetry of the problem, we choose $\delta_0=\delta$ and $\delta_1=1-\delta$, and we do all our analysis assuming $\rho(0)=1$. Thus, the filter equation is:

$$d\pi_1(t) = K^2 \pi_1(t) (1-\pi_1(t))^2 dt + K\pi_1(t) (1-\pi_1(t)) dw(t) \quad (7.3.2)$$

and we take $\pi_1(0)=1/2$ as initial condition.

Equation (7.3.2) is a diffusion confined to the interval $[0,1]$. The behavior of one-dimensional diffusions in bounded domains is well understood (see for example [Kar 81]), and the performance parameters adequate for this case - probability of false detection, mean time to detection and mean time between false alarms - can be analyzed using standard results from stochastic process theory. We have summarized the results needed from this theory in Appendix 7.A where we also establish the notation used in the sequel.

7.3.1 Boundary Classification

As discussed in Appendix 7.A, the nature of the boundaries is of crucial importance in the analysis of diffusions in bounded domains. We now compute the scale function and the speed density function for the process (7.3.2) whose drift coefficient is:

$$\mu(x) = K^2 x(1-x)^2 \quad (7.3.3)$$

and noise intensity is

$$\sigma^2(x) = K^2 x^2(1-x)^2 \quad (7.3.4)$$

The scale function is then found to be:

$$s(x) = \int^x \exp \left\{ - \int_{\eta}^x \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi \right\} d\eta = \int_{\eta}^x \frac{1}{\eta^2} d\eta = - \frac{1}{x} \quad (7.3.5)$$

and the speed density

$$m(x) = \frac{1}{\sigma^2(x) \exp\left\{-\int_0^x \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi\right\}} = \frac{1}{K^2(1-x)^2} \quad (7.3.6)$$

Notice that $m(x) = \infty$ at $x=1$ while it is finite at $x=0$. According to the interpretation given to $m(x)$ in Appendix 7.A, this means that the process moves very slowly near 1 while it moves at a certain speed when it is close to zero, suggesting that $\pi_1(t)$ tends to stay close to 1. In fact, as defined in Appendix 7.A, 1 is an attracting boundary but zero is not, because

$$S(0, x] = \lim_{a \rightarrow 0} S(x) - S(a) = \infty$$

$$S[x, 1) = \lim_{b \rightarrow 1} S(b) - S(x) = \frac{1}{x} - 1 < \infty, \quad x > 0$$

However, although 1 is an attracting boundary, it is not reached in finite time because

$$\sum(1) = \lim_{b \rightarrow 1} \int_x^b S(\eta, 1) m(\eta) d\eta = \infty$$

and therefore 1 is unattainable. Furthermore,

$$M[x, 1) = \lim_{b \rightarrow 1} \int_x^b \frac{1}{K^2(1-x)^2} = \infty$$

which implies that if $\pi_1(t)$ is started at the boundary $\pi_1(0)=1$, it does not reach the interior in finite time. The conclusion is that 1 is a *natural* boundary for $\pi_1(t)$.

As said in Appendix 7.A the nature of these boundaries implies that

$$\lim_{t \rightarrow \infty} \pi_1(t) = 1$$

with probability one, i.e., the filter eventually converges to the correct probability. It is clear that the trade-off involved in the selection of the detection level δ is between the expected time to detection and the probability of a false detection, the larger δ is the faster the detection is made but the larger is the error probability. Again, we can use the results in the Appendix 7.A to quantify this trade-off.

7.3.2 Probability of Error Versus Mean Time to Detection

Assuming $\rho(0)=1$ and $\pi_1(0)=1/2$, the probability of an erroneous detection is the probability that $\pi_1(t)$ will exit $(\delta, 1-\delta)$ through the lower and δ . It is given by:

$$1 - u_{1/2}(\delta, 1-\delta) = \frac{S(1-\delta) - S(1/2)}{S(1-\delta) - S(\delta)} = \frac{2 - \frac{1}{1-\delta}}{\frac{1}{\delta} - \frac{1}{1-\delta}} = \delta \quad (7.3.7)$$

The expected time to detection, on the other hand, turns out to be:

$$v_{1/2}(\delta, 1-\delta) = 2 \int_{1/2}^{1-\delta} \left(-\frac{1}{1-\delta} + \frac{1}{\xi} \right) \frac{d\xi}{K^2 (1-\xi)^2} \\ + 2\delta \int_{\delta}^{1/2} \left(-\frac{1}{\xi} + \frac{1}{\delta} \right) \frac{d\xi}{K^2 (1-\delta)^2}$$

and the evaluation of these integrals (see Appendix 7.B) shows that:

$$v_{1/2}(\delta, 1-\delta) = 0 \left(-\frac{\ln \delta}{K^2} \right) \quad (7.3.8)$$

Notice that the probability of a false detection does not depend directly on the signal-to-noise ratio, only the mean time to detection does. This fact seems to go against the usual view that high signal-to-noise ratio should result in low error probabilities. There is however, no contradiction. The results should be interpreted as saying that we can achieve an arbitrary small false alarm probability provided we are ready to wait a long period before making a decision. On the other hand, for a given mean time to detection (say 1), the detection level must be chosen of order $\exp\{-K^2\}$ and in this case the error probability is also of order $\exp\{-K^2\}$ which shows the usual pattern.

7.3.3 Probability of Large Deviations

To conclude our analysis of the hypothesis testing problem we study the occurrence of large deviations from the average behavior of the filter. As we said before this issue is not typically looked at in hypothesis testing problem as it has no relevance in those problems. We will find it useful, however, when contrasting the behavior of the filter for a process with rare transitions with the no-transitions case.

Suppose that $\pi_1(t)$ has reached the detection level $1-\delta$ corresponding to the correct decision $\rho(0)=1$ and consider this hitting time to be the new time origin, i.e., $\pi_1(0) = 1-\delta$. What is the probability that $\pi_1(t)$ will have a large deviation and cross the level δ before converging to 1? If such a large excursion occurs, what is the expected time until it takes place?

The probability of a large excursion is:

$$\begin{aligned} 1-u_{1-\delta}(\delta,1) &= \frac{S(1)-S(1-\delta)}{S(1)-S(\delta)} = \\ &= \frac{\frac{1}{1-\delta} - 1}{\frac{1}{\delta} - 1} = \frac{\delta^2}{(1-\delta)^2} = o(\delta^2) \end{aligned} \quad (7.3.9)$$

and, assuming that such excursion takes place, the expected exit time is given by (see eq. (7.A.9)):

$$\begin{aligned}
 v_{1-\delta}(\delta, 1) &= 2 \int_{\delta}^{1-\delta} [S(\xi) - S(\delta)] m(\xi) d\xi \\
 &= 2 \int_{\delta}^{1-\delta} \left[\frac{1}{\delta} - \frac{1}{\xi} \right] \frac{d\xi}{K^2 (1-\xi)^2} \\
 &= \frac{2}{K^2} \cdot \frac{1-2\delta}{\delta^2 (1-\delta)} + \frac{2}{K^2} \left[-2 \ln \frac{\delta}{1-\delta} + \frac{1-2\delta}{\delta(1-\delta)} \right] \\
 &= O\left(\frac{1}{\delta^2 K^2}\right) \quad \text{as } \delta \downarrow 0
 \end{aligned}$$

Notice that in the above equation a high signal-to-noise ratio seems to have an adverse effect by reducing the expected time to a large excursion. Indeed that is the case if δ is kept fixed and $K \rightarrow \infty$. The explanation is that at high signal-to-noise ratios a large confidence is posed on the measurements and as a result the measurement noise is more likely to produce large excursions over long time intervals. If, on the other hand, the detection level δ is chosen, as discussed before, to be of order $\exp\{-K^2\}$, then the mean time to a large excursion

$$v_{1-\delta}(\delta, 1) = O\left(\frac{\exp\{2K^2\}}{K^2}\right)$$

goes to ∞ as $K \rightarrow \infty$, clearly indicating the positive effect of large signal-to-noise ratios.

To conclude our analysis of the $\lambda=0$ case let us briefly summarize the basic ideas and results. We have established that the output of the filter (7.3.2) will converge towards the correct end point of the interval $[0,1]$ according to whether $\rho(0)=0$ or $\rho(0)=1$. We have quantified the trade-off between the error probability and the mean time to detection in terms of the selection of the detection level δ . For $\delta \ll 1$, the probability of an erroneous detection has been found to be of order δ and the mean time to detection of order $\ln(1/\delta)/K^2$. Large deviations after a correct detection have been shown to occur with probability of order δ^2 with a mean time between large deviations of order $1/\delta^2 K^2$. In the next section we will use these results to analyze the rare transitions case ($\lambda \ll 1$).

7.4 Detection of Rare Events

We now go back to the analysis of the filtering equations for a two state process when transitions between these two states are possible but rare. To make explicit the fact that we are interested in the filter performance in the limit as these transition rates become very small, we will assume that the process to be estimated, $\rho^\varepsilon(t)$, has transition rates $\varepsilon\lambda$ as opposed to λ . The goal is then to study

the limiting behavior of the qualitative performance measures proposed in Section 7.2 in the limit as $\epsilon \downarrow 0$.

These performance measures depend now on the signal-to-noise ratio K^2 and on the information rate $\gamma = K^2 / \epsilon \lambda$. As we will see, they clearly signal the trade-off between detection delays and false alarm probabilities and rates characteristic of detection problems. The asymptotic calculations as $\epsilon \downarrow 0$ are first carried out for a fixed value of K and afterwards a dependence on ϵ is also introduced in the signal-to-noise ratio to model different limiting situations.

7.4.1 The Effect of the Perturbation: Boundary Classification

In the presence of rare transitions, the filter equation (assuming $\rho^\epsilon(t) = 1$) becomes:

$$\begin{aligned} d\pi_1^\epsilon(t) = & \epsilon \lambda (1 - 2\pi_1^\epsilon(t)) dt \\ & + K^2 \pi_1^\epsilon(t) (1 - \pi_1^\epsilon(t))^2 dt \\ & + K \pi_1^\epsilon(t) (1 - \pi_1^\epsilon(t)) dw(t) \end{aligned}$$

and the only difference with the hypothesis testing problem (see equation (7.3.2)) is the addition of the small term $\epsilon \lambda (1 - 2\pi_1^\epsilon(t))$ to the drift coefficient.

As indicated in Figure 7.6, this small term affects the evolution of the a-posteriori probability $\pi_1(t)$ only in boundary layers close to the end points 0 and 1. As we will now see, however, the perturbation is singular and fundamentally alters the nature of these boundaries and the filter behavior.

Assume again that $\rho^\varepsilon(t)=1$. The drift coefficient is now:

$$\mu^\varepsilon(x) = \varepsilon(-\lambda x + \lambda(1-x)) + K^2 x(1-x)^2 \quad (7.4.1)$$

while the noise intensity remains unchanged:

$$\sigma(x) = Kx(1-x) \quad (7.4.2)$$

Unfortunately, the scale function and the speed measure can no longer be given in a closed form, and the computations of exit probabilities and exit times becomes quite involved. To keep the flow of ideas clear we relegate the details of these computations to several appendices at the end of the chapter.

The scale density (see Appendix 7.C) is found to be:

$$s^\varepsilon(\eta) = \frac{1}{\eta^2} \exp\left\{-\frac{2\varepsilon\lambda}{K^2} \frac{2\eta-2\eta^2-1}{\eta(\eta-1)}\right\} \quad (7.4.3)$$

Figure 7.7 shows that $s^\varepsilon(\eta)$ closely resembles the scale density for the no-transition ($\varepsilon=0$) case except in boundary layers of width $\varepsilon\lambda/K^2 = 1/\gamma$ around 0 and 1. Precisely,

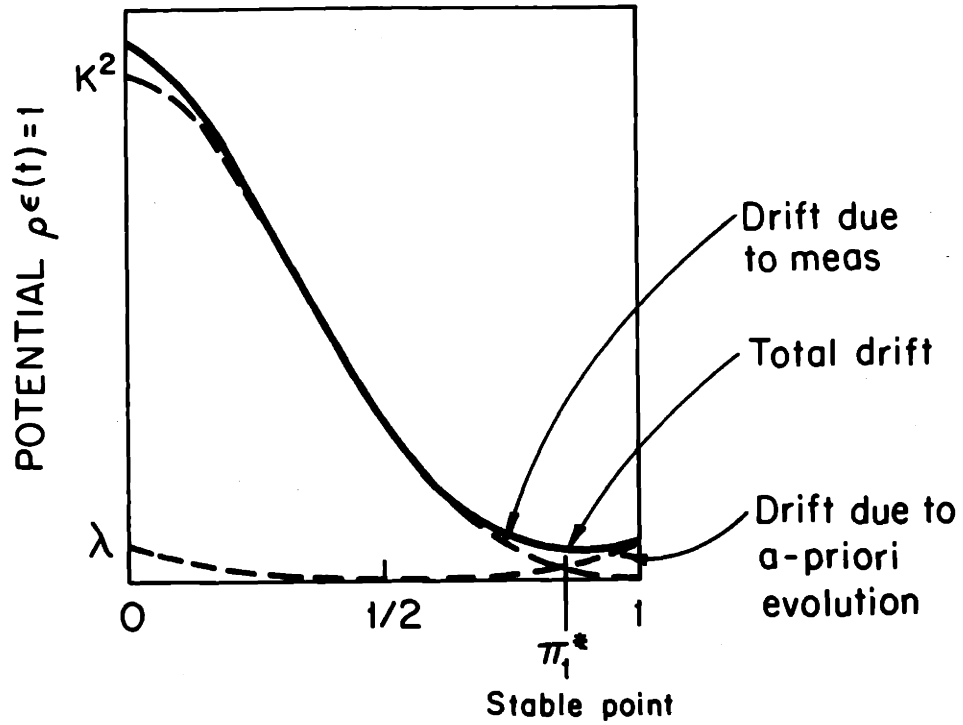


Figure 7.6a: The drift coefficient.

$$O\left(\sqrt{\epsilon\lambda/K^2}\right)$$

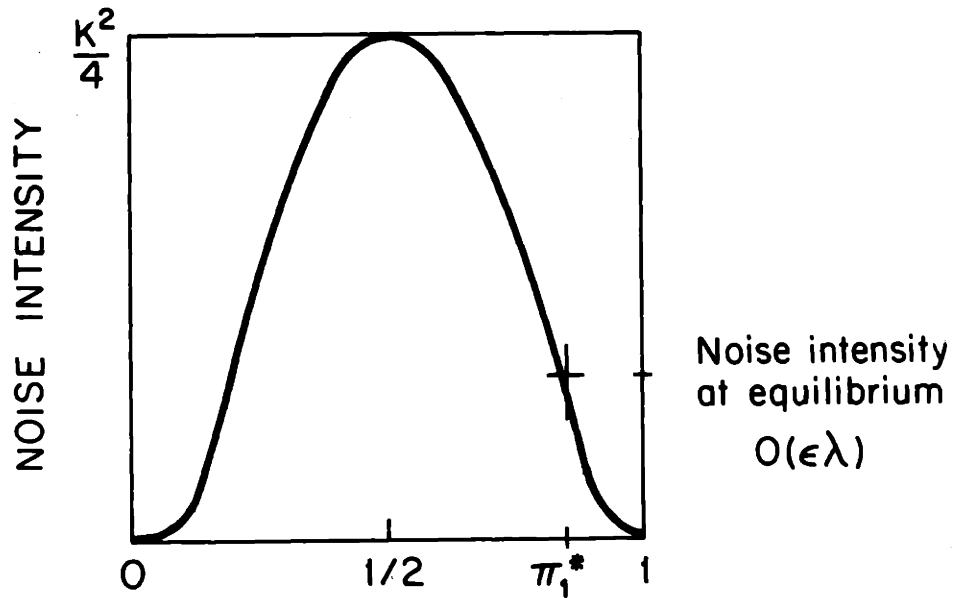
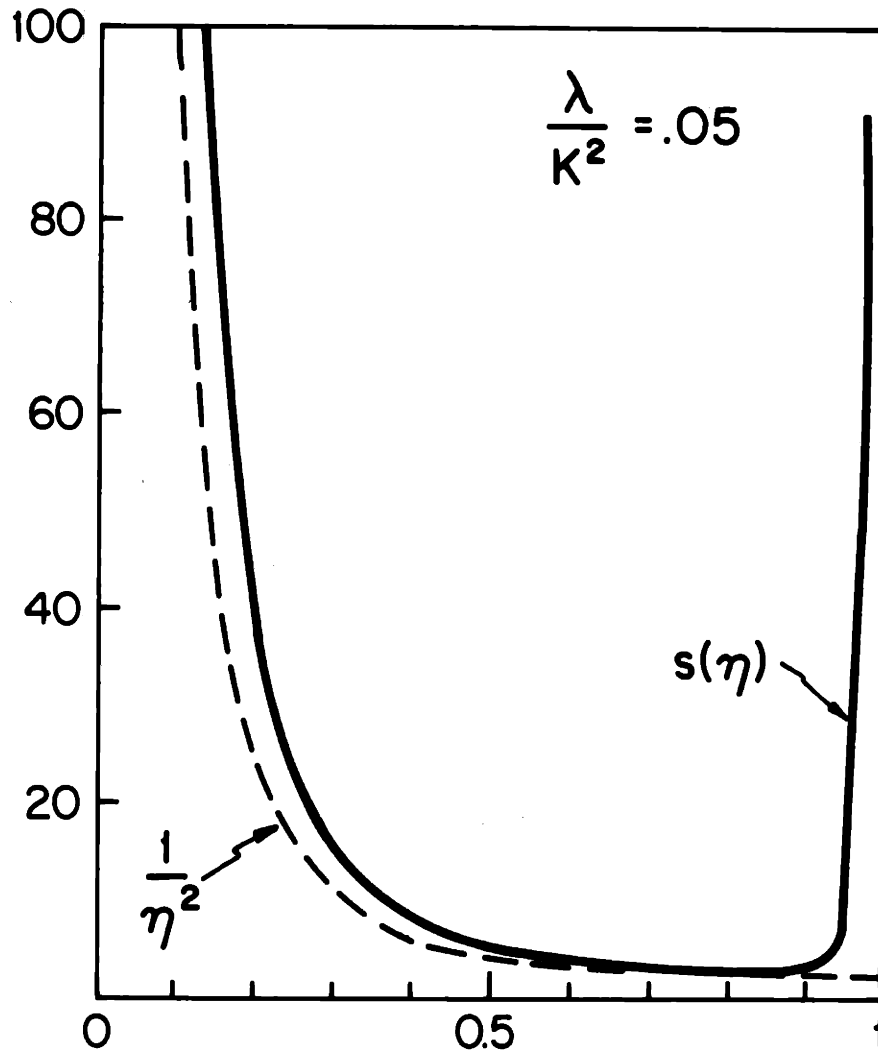


Figure 7.6b: Potential and noise intensity for the rare transitions case.



$$s(\eta) = \frac{1}{\eta^2} \exp \left\{ - \frac{2 \lambda}{K^2} \left(\frac{2\eta - 2\eta^2 - 1}{\eta(1-\eta)} \right) \right\}$$

Figure 7.7: Scale density.

$$\lim_{\epsilon \downarrow 0} \sup_{\eta \in [b/\gamma^q, 1-b/\gamma^q]} \left| s^\epsilon(\eta) - \frac{1}{\eta^2} \right| = 0 \quad (7.4.4)$$

$$\forall b > 0, 0 < q < 1$$

The speed density,

$$m^\epsilon(x) = \frac{\exp\left\{ \frac{2\epsilon\lambda}{K^2} \left(\frac{2x-2x^2-1}{x(x-1)} \right) \right\}}{K^2(1-x)^2} \quad (7.4.5)$$

also exhibits this boundary larger behavior and, in contrast with the $\epsilon=0$ case (see eq. (7.3.6)), it is now zero at the two boundaries indicating that neither of them is attracting. As we will now see, these layers radically change the nature of the boundaries no matter how small ϵ is. On the other hand, exit probabilities and mean exit times which *do not* involve states inside the boundary layers can be well approximated by the corresponding quantities computed for $\epsilon=0$.

To determine the nature of the boundaries requires that one consider in detail the behavior of $s^\epsilon(x)$ and $m^\epsilon(x)$ inside the boundary layers. It is clear that for $\forall \epsilon > 0$,

$$s^\epsilon[a, 1] = \int_a^1 s(\eta) d\eta \approx \int_a^1 \exp\left\{ \frac{2\epsilon\lambda}{K^2} \frac{1}{\eta-1} \right\} d\eta = \infty \quad (7.4.6)$$

* The approximation made is the replacement of

$$\frac{2\eta-2\eta^2-1}{\eta(\eta-1)} \quad \text{with} \quad -\frac{1}{\eta-1} \quad \text{or} \quad \frac{1}{\eta} \quad \text{for } \eta \text{ close to } 1 \text{ or } 0 \text{ respectively.}$$

and

$$s^\varepsilon(0,b] = \int_0^b s^\varepsilon(\eta) d\eta \approx \int_0^b \frac{1}{\eta^2} \exp\left\{\frac{2\varepsilon\lambda}{K^2} \frac{1}{\eta}\right\} d\eta = \infty$$

Thus, both boundaries are *non-attracting*. Intuitively, the boundaries should be *entrance*, i.e., even if we start with perfect knowledge ($\pi_1^\varepsilon(t)=0$ or 1), $\pi_1^\varepsilon(t)$ should evolve to the interior of $[0,1]$ in a finite time. Notice, on the other hand, that the boundaries are never reached from the interior.

In Appendix 7.C it is shown that

$$N^\varepsilon(1) = \int_x^1 \int_x^\xi s^\varepsilon(\eta) d\eta \cdot m^\varepsilon(\xi) d\xi \leq M \cdot 1/\varepsilon \quad (7.4.7)$$

which establishes that 1 is indeed an entrance boundary for any $\varepsilon > 0$

(the same computations can be performed to show that $N^\varepsilon(0) < \infty$).

Notice that $N^\varepsilon(1) \rightarrow \infty$ as $\varepsilon \rightarrow 0$ because, as we have seen, 1 becomes a natural boundary. The inequality (7.4.7) also indicates that starting at 1, the mean hitting time to an interior state x is at most of order $1/\varepsilon$ as $\varepsilon \downarrow 0$.

It follows from the above discussion that in the presence of rare transitions the a posteriori probability $\pi_1^\varepsilon(t)$ does not converge to the appropriate end point of $[0,1]$ although it tends to remain

most of the time very close to it. For example, assuming that *no transition occurs* in $\rho^\epsilon(t)$, the probability $\pi_1^\epsilon(t)$ settles to an invariant measure given by:

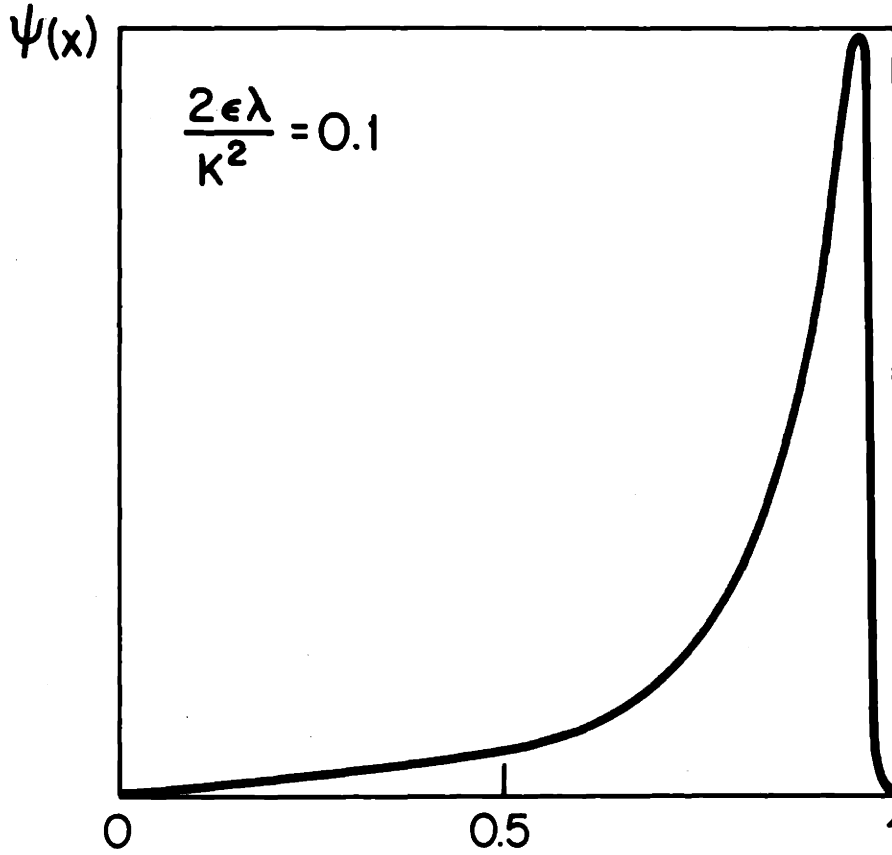
$$\psi^\epsilon(x) = \frac{\frac{1}{(1-x)^2} \exp\left\{\frac{2\epsilon\lambda}{K^2} \frac{2x-2x^2-1}{x(1-x)}\right\}}{\int_0^1 \frac{1}{(1-x)^2} \exp\left\{\frac{2\epsilon\lambda}{K^2} \frac{2x-2x^2-1}{x(1-x)}\right\} dx} \quad (7.4.8)$$

which (see Figure 7.8) is very skewed for $\frac{\lambda}{K^2}$ small around the stable point $\pi_1^* \approx 1 - \sqrt{\lambda/K^2}$. These are however heuristic notions which say little about the *dynamic* behavior of the filter as far as, for example, how should we select levels for detection purposes and how will the filter perform in terms of false alarms probabilities and detection delays. We address these issues in the following sections.

7.4.2 Mean Time Between False Alarms and Mean Time to Detection

Suppose that $\rho^\epsilon(0)=1$ and, as in Section 7.3, suppose that crossing levels δ and $1-\delta$ are chosen for detection purposes. If $\pi_1(0)=1/2$, the probability of a correct detection, i.e., the probability of hitting $1-\delta$ before δ , is (for ϵ small enough) close to the same quantity for $\epsilon=0$. That is, of order $O(1-\delta)^*$. After the first detection, however, the behavior of the filter for $\epsilon>0$ differs markedly from the limit case $\epsilon=0$.

* Notice that these estimates refer to a double limit in the following order:
 $\lim_{\delta \downarrow 0} \lim_{\epsilon \downarrow 0}$.



$$\psi^\epsilon(x) = \frac{\frac{1}{(1-x)^2} \exp\left\{\frac{2\epsilon\lambda}{\kappa^2} \frac{2x-2x^2-1}{x(1-x)}\right\}}{\int_0^1 \frac{1}{(1-x)^2} \exp\left\{\frac{2\epsilon\lambda}{\kappa^2} \frac{2x-2x^2-1}{x(1-x)}\right\} dx}$$

Figure 7.8: Steady state probability density assuming $\rho^\epsilon(t)=1$ for $t \geq 0$.

The main difference is due to the change in the nature of the boundary 1 from absorbing to entrance. As we will see, $\pi_1^\epsilon(t)$ can now have relatively frequent large excursions producing false alarms between transitions in $\rho^\epsilon(t)$. The frequency of such false alarms depends on the values of K and δ and our goal in this section is to show that by making the detection level δ a function of ϵ , it can be adjusted to keep the false alarm rate at a tolerable level without losing the ability to keep track of changes in $\rho^\epsilon(t)$.

To estimate the mean time between false alarms, we compute the mean time to crossing the level δ assuming $\pi^\epsilon(0)=1-\delta$ and that $\rho^\epsilon(t)$ remains at state 1. This quantity is given by:

$$v_{1-\delta}^\epsilon[\delta,1) = 2 \int_{\delta}^{1-\delta} \int_{\delta}^{\xi} s^\epsilon(\eta) d\eta m^\epsilon(\eta) d\xi \quad (7.4.9)$$

and because $s^\epsilon(\eta)$ and $m^\epsilon(\xi)$ converge uniformly on $[\delta,1-\delta]$ as $\epsilon \downarrow 0$ to the corresponding quantities for $\epsilon=0$, it follows from equation (7.3.10) that the mean time between false alarms satisfies

$$\lim_{\epsilon \downarrow 0} v_{1-\delta}^\epsilon[\delta,1) = 0 \left(\frac{1}{\delta^2 K^2} \right) \quad (7.4.10)$$

Thus, because the mean time between transitions in $\rho^\epsilon(t)$ is of order $1/\epsilon$, for fixed δ and K and small ϵ , a large number of false alarms (of the order of $\delta^2 K^2/\epsilon$) occurs before the correct detection of a transition in $\rho^\epsilon(t)$. Notice that as in the hypothesis testing

case a large signal-to-noise ratio seems to have an adverse effect by reducing the mean time between false alarms. Again, the same comments apply. The real positive effect of high signal-to-noise ratios will be apparent when the detection level δ is properly adjusted.

Obviously, a detection threshold $\delta(\epsilon) \downarrow 0$, as $\epsilon \downarrow 0$, is required to avoid these multiple false alarms caused by large deviations of $\pi_1^\epsilon(t)$. With a detection level very close to the boundaries, however, the detection delay is increased, pointing out again the familiar trade-off in detection problems.

To compute the detection delay we assume that at time $t=0$ $\rho^\epsilon(t)$ jumps from $\rho^\epsilon(0^-)=0$ to $\rho^\epsilon(0^+)=1$. The filter output $\pi_1^\epsilon(t)$ is assumed to be at its equilibrium point $\pi_0^* = \sqrt{\frac{\epsilon\lambda}{K^2}}$ at $t=0$ and the quantity of interest is the expected value of the time it takes for $\pi_1^\epsilon(t)$ to reach $1-\delta(\epsilon)$ assuming $\pi_1^\epsilon(0) = \pi_0^*$, and that $\rho^\epsilon(t)$ remains at 1.

The quantity of interest is then given by

$$v_{\pi_0^*}^\epsilon(0, 1-\delta(\epsilon)) = \int_{\pi_0^*}^{1-\delta(\epsilon)} \int_{\xi}^{1-\delta(\epsilon)} \frac{1}{\eta^2} \exp\left\{\frac{2\epsilon\lambda}{K^2} \frac{1}{1-\eta}\right\} \cdot \frac{\exp\left\{-\frac{2\epsilon\lambda}{K^2} \frac{1}{1-\xi}\right\}}{K^2(1-\xi)^2} d\eta d\xi \quad (7.4.11)$$

and its asymptotic evaluation depends on the order in ϵ of the threshold $\delta(\epsilon)$. Based on the shape of the drift $\mu(x)$ (see Fig. 7.6)

we choose a detection level close to the stable point of the diffusion, i.e.

$$\delta(\lambda) = O\left(\sqrt{\frac{\epsilon\lambda}{K^2}}\right) \quad (7.4.12)$$

which gives:

$$\exp\left\{\frac{2\epsilon\lambda}{K^2} \frac{1}{1-\xi}\right\} = O(1) \quad \text{as } \epsilon \downarrow 0$$

uniformly over $[\pi_0^*, 1-\delta(\epsilon)]$. Using this approximation, the detection delay can be computed as for the hypothesis testing problem (see Appendix 7.B) giving:

$$v_{\pi_0^*}(0, 1-\delta(\epsilon)) \approx \int_{\pi_0^*}^{1-\delta(\epsilon)} \int_{\xi}^{1-\delta(\epsilon)} \frac{1}{\eta} d\eta \frac{1}{K^2(1-\xi)} d\xi = O\left(-\frac{\ln \delta(\epsilon)}{K^2}\right) \quad (7.4.13)$$

Thus, if a threshold $\delta(\epsilon) = O\left(\sqrt{\epsilon\lambda/K^2}\right)$ is chosen, the detection delay goes to infinity as $\epsilon \downarrow 0$ at a rate $-(\log \delta(\epsilon))/K^2$.

Notice, however, that this delay is still small compared to the interval between transitions in $\rho^\epsilon(t)$ which is of order $1/\epsilon$. Thus, when the delay is normalized to the time scale at which $\rho^\epsilon(t)$ jumps, the delay goes to zero as $\epsilon \downarrow 0$ and therefore the filter is able to track $\rho^\epsilon(t)$ in the sense that it is correct "most of the time".

Let us now compute the mean time between false alarms for $\delta(\epsilon)$ as in (7.4.12). Assuming $\rho^\epsilon(0)=1$, no transitions in $\rho^\epsilon(t)$, and $\pi_1^\epsilon(0) = 1-\delta(\epsilon)$, we want to estimate the mean hitting time to $\delta(\epsilon)$ which is given by:

$$v_{1-\delta(\epsilon)}[\delta(\epsilon), 1] = \int_{\delta(\epsilon)}^{1-\delta(\epsilon)} \int_{\delta(\epsilon)}^{\xi} s^\epsilon(\eta) d\eta m^\epsilon(\xi) d\xi \quad (7.4.14)$$

As shown in Appendix 7.D, the asymptotic behavior of (7.4.14) is found to be: *

$$v_{1-\delta(\epsilon)}[\delta(\epsilon), 1] = o\left(\frac{1}{K^2 \delta(\epsilon)^2}\right) = o(1/\epsilon) \quad (7.4.15)$$

We now have all the ingredients to evaluate the trade-off between false alarm and detection delays. By choosing a detection threshold $\delta(\epsilon)$ of order $\epsilon\lambda/K^2$, the mean time between false alarms

* It is interesting to note the difference in order of magnitude of this escape time with the exponential order $\exp\{-1/\epsilon^2\}$ obtained by Ventcel and Friedlin [Ven 70] for diffusions with small noise intensity (or order ϵ). Although the noise intensity around the stable point π_1^* is of order ϵ in the problem we consider, it is not uniformly small over the interval $[0,1]$ and once $\pi_1^\epsilon(t)$ exits a neighborhood of π_1^* the noise intensity rapidly increases making large excursions more likely.

is of order $1/\epsilon$, independent of the signal-to-noise ratio K^2 .

This mean time between false alarms is comparable to the length of the time interval between transitions to be detected. Thus, there is a certain probability of a false alarm before a transition in $\rho^\epsilon(t)$ occurs and is detected, but this probability neither becomes zero nor one as $\epsilon \downarrow 0$. The probability of more than one false alarm, however, converges to zero as $\epsilon \downarrow 0$. The precise value of the false alarm probability depends on the proportionality constant used in $\delta(\epsilon)$ and on parameters like λ and K^2 . For given parameters, its determination requires the numerical evaluation of integrals like (7.4.14).

Multiple false alarms have been avoided at the cost of a delay $\tau(\epsilon)$ in detecting transitions of the order of

$$\tau(\epsilon) = 0 \left(- \frac{\ln \sqrt{\frac{\lambda \epsilon}{K^2}}}{K^2} \right)$$

The detector however is right "most of the time" in the sense that the ratio between length of time in which the detector gives a false reading (delay in detection plus false alarms), and the interval of time with correct detection converges to zero as $\epsilon \downarrow 0$:

$$\lim_{\epsilon \downarrow 0} \frac{\ln \sqrt{K^2/\lambda \epsilon} / K^2}{1/\epsilon \lambda - \ln \sqrt{K^2/\lambda \epsilon} / K^2} = 0 \quad (7.4.16)$$

Notice that in the numerator of (7.4.15) we have used the fact that a false alarm produces an erroneous reading during an interval of time of the same order of magnitude as the detection delay.

From the analysis presented above we conclude that when trying to detect rare transitions the thresholds used in the detector must be chosen to be commensurate with the rarity of the events to be detected. Otherwise, a probability of false alarm very close to 1 can be expected. This choice for the threshold increases the detection delay but this delay is still negligible when compared with the length of the time interval between transitions. It is in this sense that we can think of high information rates (of the order of $\gamma = K^2/\epsilon\lambda$ in this case) as resulting in near perfect detection. It must be noticed, however, that this good detector performance does not exclude sporadic, short-duration false alarms.

Instead of thinking in terms of rare transitions, we can interpret a high information rate $\gamma = K^2/\epsilon\lambda$ as resulting from a problem with transitions of rate λ and high signal-to-noise ratio $K^2(\epsilon) = K^2/\epsilon$. If the detection level δ is chosen independent of ϵ , the mean time between false alarms (see eq. (7.4.10)) turns out to be of order $\epsilon/\delta^2 K^2$. Thus, as in the low rate case, a large number of false alarms occur before a transition is detected (the interpretation is that high signal-to-noise ratios open up the bandwidth

of the filter and, as a result, the measurement noise is more likely to produce large excursions in the output of the filter). Following the same reasoning as before, we should choose a detection level $\delta(\epsilon) = O(\sqrt{\epsilon\lambda/K^2})$. In this way, we avoid multiple false alarms and the detection delay, of order $-\epsilon \log \delta(\epsilon)/K^2$ (see eq. (7.4.13)), is, as before, negligible when compared to the interval between transitions. The detector is again correct "most of the time." In fact, this case in which $\lambda(\epsilon)=\lambda$ and $K^2(\epsilon) = K^2/\epsilon$ is obtained simply by changing the time scale from t to ϵt in the problem formulated earlier with $\lambda(\epsilon) = \epsilon\lambda$ and $K^2(\epsilon) = K^2$.

To close this section let us briefly summarize the main ideas and results presented. We have shown that in the detection of rare events there is a fundamental trade-off between detection delays and the probability and frequency of false alarms. We have established that by choosing the detection levels commensurate with the rarity of the transitions to be detected, this trade-off can be resolved in such a way that the mean time between false alarms is of the same order of magnitude as the mean time between transitions to be detected, and, at the same time detection delays are negligible when compared to this mean time between transitions. This has lead us to introduce the notion of measuring a filter or detector performance in terms of whether or not it is right "most of the time." We will use this notion in the section that follow to argue for filters with a hierarchical structure.

7.5 Filtering with Aggregated Measurements

In this section we propose a filter decomposition based on aggregation ideas for the example discussed in Section 7.1 (see Fig. 7.9). We first analyze in Section 7.5.1 the case in which only information about the slow transitions between groups of states is available. For this case, a reduced-order filter is proposed to estimate the aggregated probabilities and the results in Section 7.4 are used to assess the performance of such filter.

In Section 7.5.2 we consider the case with separate channels of information for the slow and fast transitions and a hierarchical filter is proposed. The discussion in this latter section is heuristic and qualitative and it is mainly intended to suggest ideas for the future work in this area.

7.5.1 Filtering with Aggregated Measurements

Consider the process $\rho^E(t) \in \{1,2,3,4\}$ shown in Figure 7.9(a) and suppose that we observe

$$dy(t) = h(\rho^E(t))dt + bdw(t) \quad (7.5.1)$$

with

$$h(1) = h(2) \stackrel{\Delta}{=} \alpha_1$$

and

$$h(3) = h(4) \stackrel{\Delta}{=} \alpha_2$$

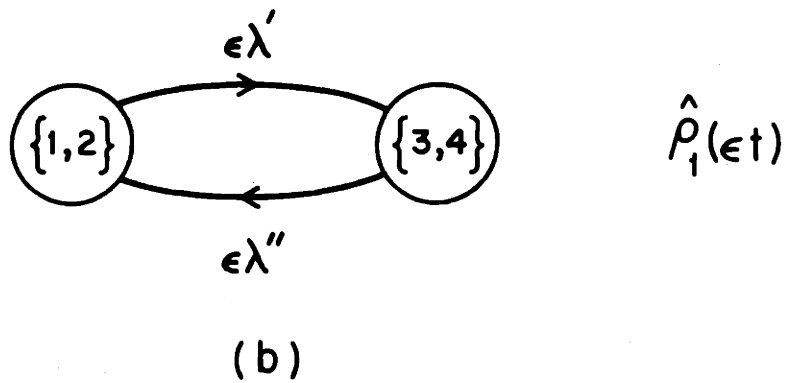
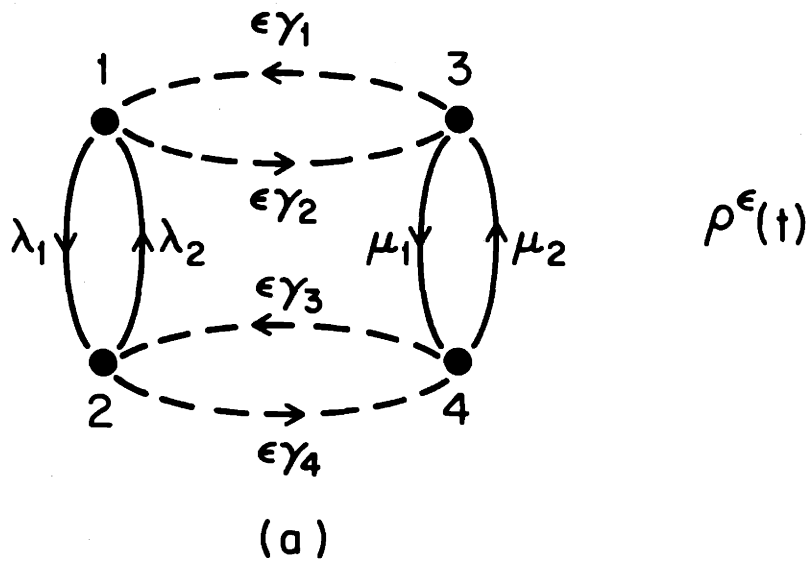


Figure 7.9: The process $\rho^\epsilon(t)$ considered in Section 7.5.1 and its aggregate approximation $\hat{\rho}_1^\epsilon(t)$.

These observations contain no information at all about the fast transitions in $\rho^\epsilon(t)$ but they carry significant information about the slow transitions (with signal-to-noise ratio $K^2 = (\alpha_1 - \alpha_2)^2 / b^2$ and information rate of order K^2/ϵ). It thus seems reasonable to expect that the only meaningful information about the evolution of $\rho^\epsilon(t)$ that can be extracted from $y(t)$ refers to transitions between $\{1,2\}$ and $\{3,4\}$ which occur only at the slow time scale t/ϵ .

It follows from our results in Chapter V that as far as slow transitions in $\rho^\epsilon(t)$ are concerned the model $\rho^\epsilon(t)$ shown in Figure 7.9(b) is a good approximation for ϵ small. The question thus arises as to whether a filter constructed using $\hat{\rho}^\epsilon(t)$ as a model to estimate the aggregated probabilities.

$$\Pr\{\rho^\epsilon(t) \in \{1,2\} \mid y(\tau), 0 \leq \tau \leq t\}$$

$$\Pr\{\rho^\epsilon(t) \in \{3,4\} \mid y(\tau), 0 \leq \tau \leq t\}$$

will have a performance close to that of the optimal filter that uses the detailed model of $\rho^\epsilon(t)$. The analysis of the optimum filter equations that we now develop strongly suggests that it is indeed the case.

The optimum filter equations written as a diffusion driven by the observation noise are:

$$\begin{aligned}
 d\pi(t) = & (A+\epsilon B)^T \pi(t) dt \\
 & + \frac{1}{2} g(\pi(t)) [h(\rho^\epsilon(t)) - h_{av}(\pi(t))] dt \\
 & + \frac{1}{b} g(\pi(t)) dw(t)
 \end{aligned} \tag{7.5.2}$$

where $\pi(t)$ and $g(\pi(t))$ are vectors with components:

$$(\pi(t))_i = \pi_i(t) = \Pr\{\rho^\epsilon(t)=i | y(\tau), 0 \leq \tau \leq t\} \tag{7.5.3}$$

$$(g(\pi(t)))_i = \pi_i(t) (h(i) - h_{av}(\pi(t))) \tag{7.5.4}$$

$A(\epsilon) = A + \epsilon B$ is the matrix of transition rates of $\rho^\epsilon(t)$, and

$$h_{av}(\pi(t)) = \sum_{i=1}^4 h(i) \pi_i(t) \tag{7.5.5}$$

Define the vector of aggregated probabilities

$$\bar{\pi}(t) \triangleq [\bar{\pi}_1(t), \bar{\pi}_2(t)]^T$$

as follows:

$$\bar{\pi}_1(t) = \pi_1(t) + \pi_2(t)$$

$$\bar{\pi}_2(t) = \pi_3(t) + \pi_4(t)$$

and use the aggregation matrix

$$V \stackrel{\Delta}{=} \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

to write

$$\bar{\pi}(t) = V^T \pi(t) \tag{7.5.6}$$

For further reference it is convenient to write a stochastic differential equation for $\bar{\pi}(t)$ obtained by using (7.5.6) into (7.5.2) as follows:

$$\begin{aligned} d\bar{\pi}(t) &= V^T (A + \epsilon B)^T \pi(t) dt \\ &+ \frac{1}{2} V^T g(\pi(t)) [h(\rho^\epsilon(t)) - h_{av}(\bar{\pi}(t))] dt \\ &+ \frac{1}{b} V^T h(\pi(t)) dw(t) \end{aligned} \tag{7.5.7}$$

This equation can be simplified by noting that

$$V^T A^T = 0 \tag{7.5.8}$$

Also, because the structure of the measurements is compatible with the aggregation (7.5.6) we have:

$$\begin{aligned} h_{av}(\pi(t)) &= \sum_{i=1}^4 h(i) \pi_i(t) \\ &= \alpha_1 \bar{\pi}_1(t) + \alpha_2 \bar{\pi}_2(t) = h_{av}(\bar{\pi}(t)) \end{aligned} \tag{7.5.9}$$

and

$$\begin{aligned} (V^T g(\pi(t)))_1 &= \pi_1(t) (\alpha_1 - h_{av}(\bar{\pi}(t))) + \pi_2(t) (\alpha_1 - h_{av}(\bar{\pi}(t))) \\ &= \bar{\pi}_1(t) (\alpha_1 - h_{av}(\bar{\pi}(t))) \end{aligned}$$

Similarly,

$$(V^T g(\pi(t)))_2 = \bar{\pi}_2(t) (\alpha_2 - h_{av}(\bar{\pi}(t)))$$

and therefore

$$V^T g(\pi(t)) = g(\bar{\pi}(t)) \tag{7.5.10}$$

Using (7.5.8)-(7.5.10) in (7.5.7) we arrive at:

$$\begin{aligned} d\bar{\pi}(t) &= \varepsilon V^T B \pi(t) dt \\ &+ \frac{1}{b^2} g(\bar{\pi}(t)) [h(\rho^\varepsilon(t)) - h_{av}(\bar{\pi}(t))] dt \\ &+ \frac{1}{b} g(\bar{\pi}(t)) dw(t) \end{aligned} \tag{7.5.11}$$

Equation (7.5.11) is the equation satisfied by the aggregate probabilities $\bar{\pi}(t)$ computed from the output of the optimum filter that uses the detailed model of $\rho^\varepsilon(t)$. Notice that if it were not for the order ε term in the drift coefficient in (7.5.11), we would be able to construct a one dimensional *optimal* filter to compute the a-posteriori aggregate probabilities (instead of the three dimensional filter (7.5.2) followed by the additions specified by

(7.5.6)). This feature is a direct consequence of the fact that the aggregated structure of the measurements coincides with the partition determined by the time scale structure of the process to be estimated. Notice also that the two terms in right-hand side of (7.5.11) due to the presence of measurements can be written entirely in terms of the aggregated probabilities $\bar{\pi}(t)$. This is a reflection of the fact that the measurements carry no information *at all* about fast transitions. The only term in the filter involving the vector of detailed probabilities $\pi(t)$ is the one corresponding to the a-priori evolution of these probabilities.

Equation (7.5.11) that computes the a-posteriori aggregated probabilities is coupled through $\pi(t)$ to equation (7.5.2). The computation of $\bar{\pi}(t)$ by means of an uncoupled equation involves an approximation that we now discuss.

Suppose that instead of $\rho^\varepsilon(t)$ we use the aggregated model $\hat{\rho}^\varepsilon(t)$ as the process to be estimated. Given the structure of the observation function $h(\cdot)$, the measurement (7.5.1) can also be written as

$$dy(t) = h(\hat{\rho}^\varepsilon(t))dt + b dw(t) \tag{7.5.12}$$

and therefore the optimum filter, *assuming that the approximate model is valid* is given by (in its realizable form):

$$\begin{aligned} d\tilde{\pi}(t) = & \epsilon V^T B^T U^T \tilde{\pi}(t) \\ & + \frac{1}{b} g(\tilde{\pi}(t)) [dy(t) - f_{av}(\tilde{\pi}(t)) dt] \end{aligned} \quad (7.5.13)$$

where $\hat{\epsilon A} = \epsilon U B V$ is the matrix of transition rates of $\hat{\rho}^\epsilon(t)$ written in terms of the aggregation matrices introduced in Chapter V.

Writing this filter as a diffusion driven by the measurement noise we get:

$$\begin{aligned} d\tilde{\pi}(t) = & \epsilon V^T B^T U^T \tilde{\pi}(t) dt \\ & + \frac{1}{2} \frac{g(\tilde{\pi}(t))}{b} [h(\rho^\epsilon(t)) - h_{av}(\tilde{\pi}(t))] dt \\ & + \frac{1}{b} g(\tilde{\pi}(t)) dw(t) \end{aligned} \quad (7.5.14)$$

The filter (7.5.14) derived using the aggregated model for $\rho^\epsilon(t)$ would coincide with the optimum filter (7.5.11) for the aggregated probabilities, if the following approximation is made:

$$\pi(t) \approx U^T \bar{\pi}(t) \quad (7.5.15)$$

This approximation is readily interpreted. In effect, it corresponds to the following:

$$\pi_1(t) \approx \frac{\lambda_2}{\lambda_1 + \lambda_2} \bar{\pi}_1(t)$$

$$\pi_2(t) \approx \frac{\lambda_1}{\lambda_1 + \lambda_2} \bar{\pi}_1(t)$$

$$\pi_3(t) \approx \frac{\mu_2}{\mu_1 + \mu_2} \bar{\pi}_2(t)$$

$$\pi_4(t) \approx \frac{\mu_1}{\mu_1 + \mu_2} \bar{\pi}_2(t)$$

That is, the a-posteriori probability of being in a given state of the detailed model is approximated by the a-posteriori probability of being in the corresponding aggregate state times the ergodic probability of the detailed state achieved due to the fast transitions inside the aggregate state.

Given that the observations do not contain any information about transitions inside the aggregate state the approximation must involve little error. To assess the loss in performance, notice that, as we have seen in Section 7.4, the ε -order terms in the filter drift are relevant only when the probability vector $\pi(t)$ is very close to the boundaries of the unit simplex where it evolves. In other words, the approximation (7.5.15) will not be accurate at times when we know that $\rho^\varepsilon(t)$ is in one of the detailed states with high probability. This may happen during an interval around $t=0$ if we start the filtering with such knowledge, or immediately following a rare transition in $\rho^\varepsilon(t)$ if the measurements $y(t)$ provide enough information to rapidly detect such a transition. In this case, detecting a transition between aggregate states may also provide some information about which of the detailed states is occupied following the rare transitions. For example, suppose that in Figure 7.9 $\gamma_3=0$. Immediately following a detected transition $\{1,2\} \rightarrow \{3,4\}$, the optimum filter will increase $\pi_3(t)$ relative to $\pi_4(t)$ because $\{3,4\}$ can only be entered through state 3. The probabilities $\pi_3(t)$ and $\pi_4(t)$ will then evolve at a fast

time scale, towards their steady state values $\mu_2/(\mu_1+\mu_2)$ and $\mu_1/(\mu_1+\mu_2)$ respectively which are precisely the values assigned to them by the suboptimal aggregate filter.

This difference between the optimal and suboptimal filter can only have, in general, minor effects on the filter performance for two reasons. First, in general, the entrance probabilities will not be concentrated in one state as in the example discussed above and therefore an aggregate transition will convey little information about the detailed states occupied immediately after this transition. Second, as we have shown in Section 7.4, there is a certain delay associated with the detection of rare transitions. This delay, although short in comparison with the interval between aggregate transitions, it is quite long when compared to the rate at which the fast transitions occur. Thus, except in extreme cases with a very high signal-to-noise ratio at the aggregate level, when the optimal filter has decided that an aggregated transition has taken place the detailed information will have disappeared with the result that the ergodic probabilities are essentially the best estimate for the detailed states. This is precisely what the suboptimal filter does all the time.

This discussion leads us to the following:

Conjecture 7.1

The diffusion $\bar{\pi}(t)$ specified by equations (7.5.2) and (7.5.11) converges weakly to $\tilde{\pi}(t)$ in (7.5.14) as $\epsilon \downarrow 0$, except perhaps at some order one intervals around the slow jumps of $\rho^\epsilon(t)$. \square

A good place to start looking for techniques to establish this result or a similar one is in the work of Papanicolau and co-workers, especially [Pap 76]. Notice that even though for clarity we have derived the aggregate filter for the example in Figure 7.7a, the same results are valid for an arbitrary n-dimensional process $\eta^\varepsilon(t)$. The above conjecture refers to the general case.

As an alternative to trying to prove the near-optimality (in the sense stated above) of the aggregated scheme, we can use the results derived in Section 7.4 to estimate the performance of the suboptimal, aggregate filter (7.5.14). We can think of (7.5.14) as a filter to estimate transitions of the two state Markov process $\hat{\rho}^\varepsilon(t)$ in Figure 7.9 (b) which is precisely the problem studied in Section 7.4. The approximation involved in this analysis is the fact that changes in $h(\rho^\varepsilon(t))$ are not exactly markovian and only in the limit as $\varepsilon \rightarrow 0$ is $h(\hat{\rho}^\varepsilon(t))$ a good model for $h(\rho^\varepsilon(t))$. Notice, however, that in the computation of performance measures in Section 7.4, we have not used *at all* the markovian character of the slow transitions. Most of these measures depend only on the behavior of the filter diffusion *for a fixed value of the process* $\rho^\varepsilon(t)$ and only the *order of magnitude* of the interval between the slow transitions enters in those performance measures. Thus, our results in Section 7.4 directly apply to the aggregate suboptimal filter $\tilde{\pi}(t)$ and therefore,

properly choosing crossing levels $\delta_0 = \delta(\epsilon)$ and $\delta_1 = 1 - \delta(\epsilon)$, we can construct a detector using the output of the aggregate filter (7.5.14) which gives the correct reading of $h(\rho^\epsilon(t))$ "most of the time". The length of intervals when the reading is erroneous converges to zero as $\epsilon \downarrow 0$ when measured in the time scale at which changes in $h(\rho^\epsilon(t))$ occur, i.e., at time scale t/ϵ .

This almost perfect detection of the rare transitions provides the justification for a hierarchical decomposition of filters that we develop in the next section where, in addition to aggregate measurements, we also introduce a separate channel of information about the fast transitions.

To close this section we state two more conjectures about the near-optimality of the aggregate filter.

Conjecture 7.2

The orders of magnitude of the quantities involved in the trade-off between detection delays and frequency and probability of false alarms for the optimal filter $\bar{\pi}(t)$ are the same as these of the corresponding quantities for the aggregate filter $\tilde{\pi}(t)$. Thus, our results in Section 7.4 also apply to the optimum filter $\bar{\pi}(t)$, and in this sense our sub-optimal filter is essentially as good as the optimal one. \square

This conjecture is based on our discussion about the negligible amount of additional information used by the optimal filter. The following conjecture refers to the occurrence of large deviations resulting in false alarms and it is based on the fact that the same noise (on a sample function basis) drives equation (7.5.11), i.e., the optimal filter and equation (7.5.14), the aggregate filter.

Conjecture 7.3

False alarms in detectors based on $\bar{\pi}(t)$ and $\tilde{\pi}(t)$ respectively, occur almost simultaneously. That is, conditioned on a large excursion of $\bar{\pi}(t)$ at time τ , the probability that $\tilde{\pi}(t)$ will not experience also a large excursion on $[\tau-\delta, \tau+\delta]$ goes to zero as $\epsilon \rightarrow 0$ for an arbitrary but fixed $\delta > 0$.

Notice that Conjecture 7.3 is *very* strong. Together with Conjecture 7.2 it states the following: not only does the suboptimal filter perform as well as the optimal filter in terms of the proportion of time they are correct, but if one of them has a false alarm, the other one will at the same time.

Both conjectures would probably be corollaries of the weak convergence result stated as Conjecture 7.1 if this result can be established.

7.5.2 Filtering with Aggregated and Decentralized Measurements

Consider again the process $\rho^E(t) \in \{1,2,3,4\}$ shown in Figure 7.9(a) but suppose now that in addition to the aggregate measurements

$$dy_1(t) = h_1(\rho^E(t))dt + b_1dw_1(t) \quad (7.5.16)$$

with $h_1(1) = h_1(2) = \alpha_1$ and $h_1(3) = h_1(4)$, a second channel of information is available with observations

$$dy_2(t) = h_2(\rho^E(t))dt + b_2dw_2(t) \quad (7.5.17)$$

where

$$h_2(1) = h_2(3) = \beta_1$$

$$h_2(2) = h_2(4) = \beta_2$$

We will refer to $y_2(t)$ as the decentralized measurement because it mainly conveys information about the fast transitions. Slow transitions have only an indirect effect on $y_2(t)$ through the change in transition rates from λ_1 and λ_2 to μ_1 and μ_2 . This information however is supplied at a slow rate because it involves observing the average behavior of $y_2(t)$ over time interval large compared to the rate of fast transitions.

Our objective is to argue that a hierarchical filter organized in two levels, each processing one of the information channels is justified as a simplified suboptimal filter.

Suppose first that the slow transitions between $\{1,2\}$ and $\{3,4\}$ are observed perfectly (for example if $b_1=0$). In this case the optimum filter has the structure shown in Figure 7.10. The perfectly observed changes in the aggregate state act as a switch changing the parameters of the fast filter. Only one filter for a two state process (i.e., a one dimensional diffusion) is required. The parameters in this filter change between λ_1, λ_2 and μ_1, μ_2 following jumps in $y_1(t)$.

If the observations $y_2(t)$ are noisy the situation is more complex and the exact filter equations do not seem to lend themselves to a hierarchical decomposition. It seems nevertheless reasonable to attempt such a decomposition by activating the switch in Figure 7.10 by the output of a detector based on $y_1(t)$, as indicated in Figure 7.11.

The detector filters the observations $y_1(t)$ using equation (7.5.14) and then it makes a zero/one decision about whether $\rho^E(t) \in \{1,2\}$ or $\{3,4\}$, by observing the crossing of the filter output $\tilde{\pi}(t)$ through detection levels chosen according to our discussion in Section 7.4. The rest of the filter works as in the case of perfectly observed aggregated transitions.

There are two sources of suboptimality in this design. First, the detector for the slow transitions will give erroneous readings due to false alarms and detection delays. As we have argued repeatedly, however, the ratio of the length of time with erroneous reading to that with correct readings goes to zero as $\epsilon \downarrow 0$. The effect of these

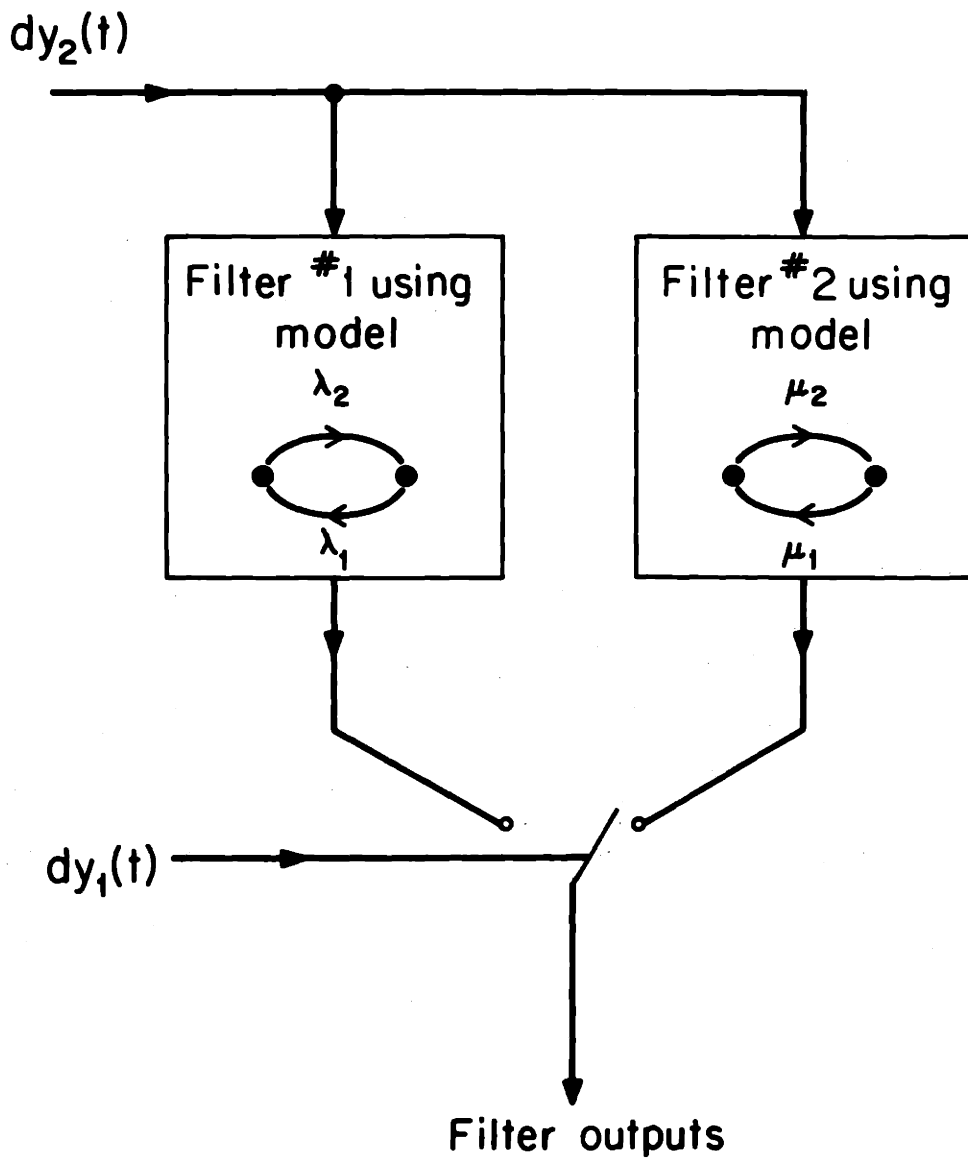


Figure 7.10: Optimum filter with perfect aggregate observations.

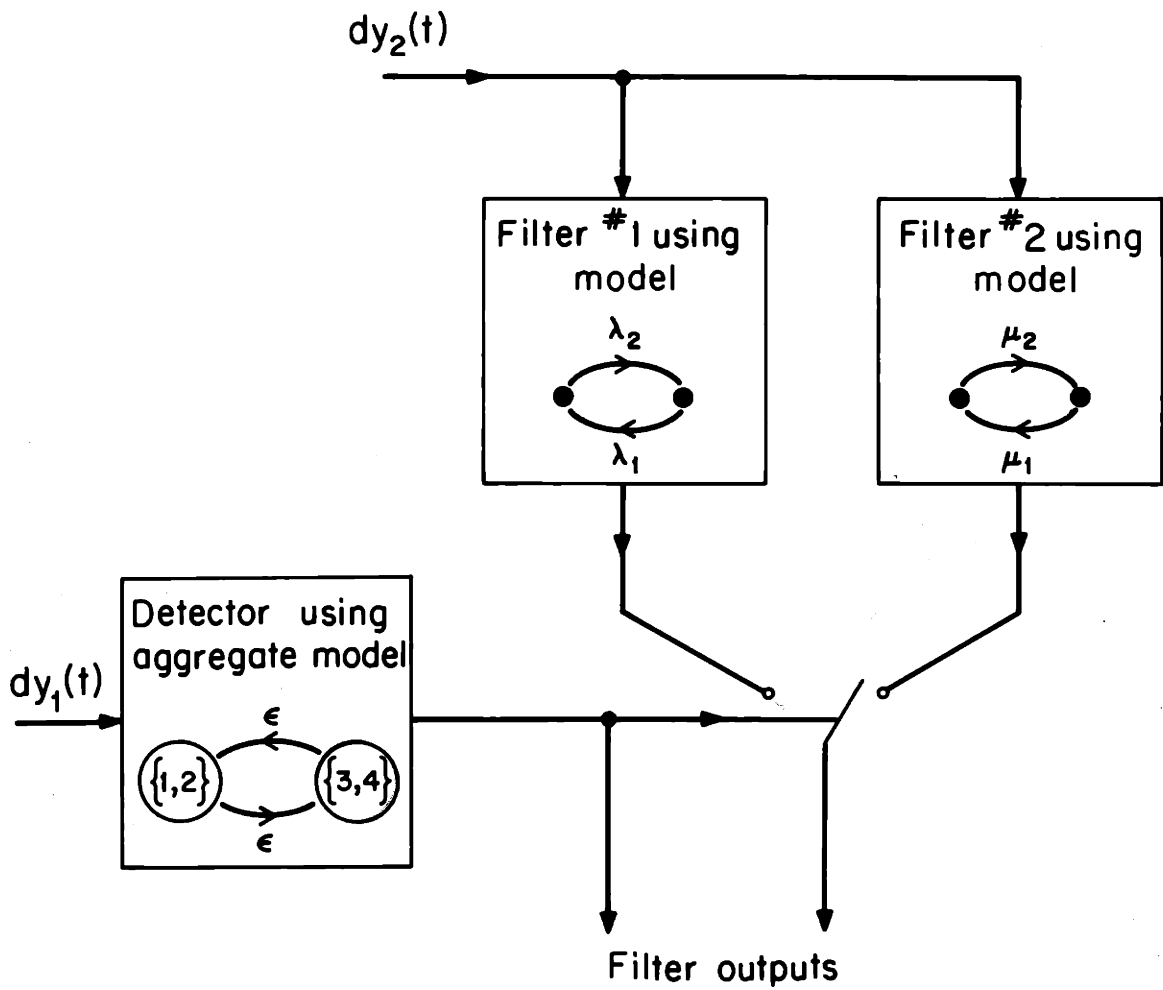


Figure 7.11: Hierarchical filter.

errors is the use of the wrong set of parameters in the fast filter and its seriousness will depend on the signal-to-noise ratio in the measurement $y_2(t)$. If this ratio is high the a-priori description of the process does not influence much the filter's evolution. In any case, the detector will give a correct reading most of the time.

The second source of suboptimality is the fact that the structure proposed in Fig. 7.11 does not make use of any information about the slow transitions supplied by the decentralized measurements $y_2(t)$. If the rates λ_1 and λ_2 differ from μ_1 and μ_2 respectively, there is some information in $y_2(t)$ about whether $\rho^\varepsilon(t) \in \{1,2\}$ or $\rho^\varepsilon(t) \in \{3,4\}$. This information is provided, however, at a slow rate because to extract it requires an averaging of $y_2(t)$ over a time interval large compared to the rate at which the fast transitions in $\rho^\varepsilon(t)$ occur. On the other hand, and unless the signal-to-noise ratio in $y_1(t)$ is zero, information about slow transitions is provided by $y_1(t)$ at a much higher rate. Our analysis thus suggests that except in extreme cases this indirect source of information can be discarded without much loss in performance.

To validate the above discussion we suggest the following tasks. Assuming that the high level aggregate filter gives a correct reading, establish the following:

Conjecture 7.4

During the interval of time when the high level filter gives a correct reading, the hierarchically decomposed filter is near optimal, i.e., the mean square difference between the output of the optimal filter and that of the aggregate filter converges to zero as $\epsilon \downarrow 0$. □

If this conjecture is proved, then it will mean that the hierarchical filter is near optimal "most of the time." To evaluate the loss of performance due to an erroneous detection at the high level filter (i.e., during detection delays are false alarms, it will probably be useful to start analyzing the loss in performance due to an erroneous model. Suppose that a filter for a two state process as that in Figure 7.12a is built, but using as a model the process in Figure 7.12b. The loss in performance clearly depends on two parameters: the difference between the rates λ and μ and the signal-to-noise ratio of the available observations. For high signal-to-noise ratios and small values of $|\lambda - \mu|$, the performance loss should be small. Estimate the asymptotic order of magnitude of the performance loss when the high level filter gives an erroneous reading for high signal-to-noise ratios in the decentralized observations and/or small differences in the fast transition rates in different aggregate states.

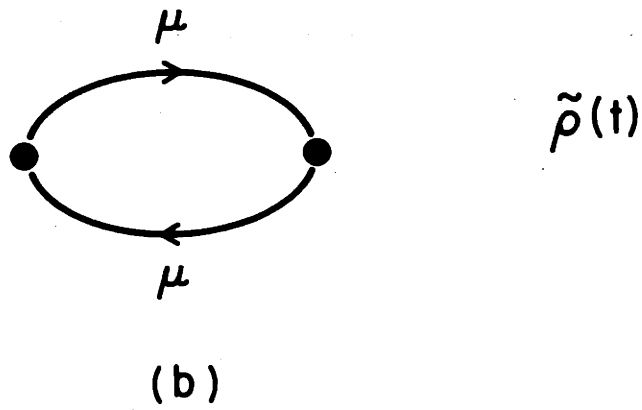
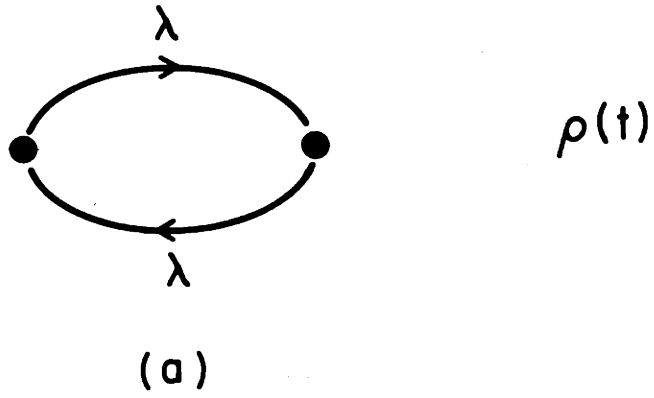


Figure 7.12(a): True process and (b) erroneous model.

In this connection we have the following conjecture:

Conjecture 7.5

The hierarchical decomposition is near optimal *at all times* in the limit as the high level transitions become increasingly rare and the signal to noise ratio of the detailed, decentralized measurements becomes larger (take, for example, $(\beta_1 - \beta_2)^2 / b_2^2 = K_2^2 / \epsilon$).

Based on the above discussion we believe that the use of hierarchical structures in filters for singularly perturbed FSMPS's is a valid engineering approach. The clarification of the circumstances under which it is a near-optimal strategy, and a rigorous theoretical basis for it still require, however, substantial research effort. It is our hope that other researchers will find the problem interesting and challenging, and that they find the ideas and insights provided here useful.

APPENDIX 7.A One-Dimensional Diffusions in Bounded Intervals

We summarize here the notation and basic results used in the analysis of filtering equations carried out in this chapter. We mostly follow [Kar 81].

Consider the stochastic differential equation:

$$dx(t) = \mu(x(t))dt + \sigma(x(t))dw(t) \tag{7.A.1}$$

in the interval $I=[\ell,r]^*$. We assume that the infinitesimal drift $\mu(x)$ is continuous in (ℓ,r) and that the noise coefficient $\sigma(x)$ satisfies $\sigma^2(x) \geq 0$ for $x \in (\ell,r)$. We are interested here in cases where due to the behavior of $\mu(x)$ and $\sigma(x)$ near the end points of I , the process $x(t)$ is "naturally" confined to evolve in I . Most of what follows is devoted to analyzing the behavior of $x(t)$ near the boundaries.

7.A.1 Basic Definitions

Hitting times of points and sets play a fundamental role in the analysis of diffusions. We define the *hitting time* of $x(t)$ to the level b starting at $x(0)=x$, by

$$T_x(b) = \begin{cases} \infty & \text{if } x(t) \neq b \text{ for } 0 \leq t < \infty \\ \inf\{t \geq 0; x(t)=b\} & \text{otherwise} \end{cases}$$

* For the development that follows we could allow $\ell=-\infty$ and/or $r=\infty$ but for our purposes it is enough to think in terms of the bounded interval case.

and $T_x(a,b)$ for $x \in (a,b)$ is used to denote the *exit time* of the interval (a,b) :

$$T_x(a,b) \triangleq \min\{T_x(a), T_x(b)\} \quad (7.A.3)$$

Two especially important functionals are the following:

i) The probability of exit through one of the end points,

$$u_x(a,b) = \Pr\{T_x(b) < T_x(a) \mid x(0)=x\} \quad x \in (a,b) \quad (7.A.4)$$

and

ii) The mean time to exit

$$v_x(a,b) = E\{T_x(a,b) \mid x(0)=x\} \quad (7.A.5)$$

They are the solution of the differential equations:

$$\begin{aligned} \mu(x) \frac{d}{dx} + \frac{1}{2} \sigma^2(x) \frac{d^2}{dx^2} u &= 0 & x \in (a,b) & \quad (7.A.6) \\ u(a) &= 0, \quad u(b) = 1 \end{aligned}$$

$$\begin{aligned} \mu(x) \frac{dv}{dx} + \frac{1}{2} \sigma^2(x) \frac{d^2 v}{dx^2} &= -1 & x \in (a,b) & \quad (7.A.7) \\ v(a) &= v(b) = 0 \end{aligned}$$

which can be explicitly solved to give:

$$u_x(a,b) = \frac{S(x) - S(a)}{S(b) - S(a)} \quad a < x < b \quad (7.A.8)$$

$$\begin{aligned} v_x(a,b) &= 2 \left\{ u_x(a,b) \int_x^b [S(b) - S(\xi)] m(\xi) d\xi \right. \\ &\quad \left. + [1 - u_x(a,b)] \int_a^x [S(\xi) - S(a)] m(\xi) d\xi \right\} \quad (7.A.9) \end{aligned}$$

where

$$S(x) = \int_l^x \exp \left\{ - \int_l^\eta \frac{2\mu(\xi)}{\sigma^2(\xi)} d\xi \right\} d\eta \triangleq \int_l^x s(\eta) d\eta \quad (7.A.10)$$

$l < x < r$

is called the *scale function* of the process,

$$m(x) = \frac{1}{\sigma^2(x) s(x)} \quad l < x < r \quad (7.A.11)$$

is called the *speed density* and $s(\eta)$ the *scale density*. For notational convenience we will also use the *scale measure*

$$S[a,b] = S(b) - S(a) = \int_a^b s(\eta) d\eta \quad (7.A.12)$$

and the *speed measure*

$$M[a,b] = \int_a^b m(x) dx \quad (7.A.13)$$

The name of scale function for $S(x)$ derives from the fact that if we use $S(x)$ to rescale the state space $[l,r]$ by defining the process $y(t) = S(x(t))$, the hitting probabilities for $y(t)$ are proportional to distances. The process $y(t)$ is then set to be in natural or canonical scale. If a process is in natural scale, the quantity $m(x)\epsilon^2$ is the order of the expected time the process spends in the interval $(x-\epsilon, x+\epsilon)$ given $x(0)=x$ before departure thereof. This interpretation motivates the name speed density for $m(x)$.

The scale measure $S[a,b]$ is monotonic in a for a fixed b and viceversa. Therefore, we can define:

$$S(\ell,b) = \lim_{a \uparrow \ell} S[a,b] \leq \infty$$

$$S[a,r) = \lim_{b \uparrow r} S[a,b] \leq \infty$$
(7.A.14)

and similarly $M(\ell,b)$ and $M[a,r)$.

To finish this section we discuss the hitting time for a boundary point, say $T_x(r)$. For a given $x < b < r$, $T_x(b)$ is a monotonically non-decreasing function of b . It follows that we may define the random time

$$T_x(r^-) = \lim_{b \rightarrow r} T_x(b)$$

and by continuity of the sample functions, $T_x(r^-) = T_x(r)$, the hitting time to the boundary r . Notice that the hitting time $T_x(r)$ is defined even when r is not a state of the process (and in this case $T_x(r) = \infty$).

7.A.2 Boundary Classification

We now give conditions that provide a way to determine whether or not the process $x(t)$ will hit the boundaries, whether or not, it will do so in a finite time, and what is the behavior of the process if

started at the boundary points. We state the results for the right boundary but they also apply to the left boundary with the obvious modifications.

Lemma 7.A.1

i) Let $S[x,r] < \infty$ for some $x \in (\ell, r)$. Then

$$\Pr\{T_x(r) \leq T_x(a) \mid x(0)=x\} > 0$$

for all $\ell < a < x < r$

ii) Let $S[x,r] = \infty$ for some $x \in (\ell, r)$. Then

$$\Pr\{T_x(r) < T_x(a) \mid x(0)=x\} = 0$$

for all $\ell < a < x < r$.

In view of the above lemma we will say that r is an *attracting boundary* if $S[x,r] < \infty$. (This criterion applies independently of $x \in (\ell, r)$ in that if it is finite for some x , it is finite for all values of x). Notice that even though there is a positive probability that the process will "exit" (a,r) through the attracting boundary r , this boundary need not be a state of the process because the probability of reaching it in a finite time may be zero. The following lemma gives a necessary and sufficient condition for an attracting boundary to be reached in finite time.

Lemma 7.A.2

Let r be an attracting boundary and suppose $l < a < x < r$.

Then the following are equivalent statements:

- i) $\Pr\{T_x(r) < \infty \mid x(0) = x\} > 0$
- ii) $E\{T_x(a, r) \mid x(0) = x\} < \infty$
- iii) $\sum(r) \triangleq \int_x^r S[\xi, r] m(\xi) d\xi < \infty$

Based on the above result we will say that an attracting boundary is *attainable* if $\sum(r) < \infty$, otherwise we will say that it is *unattainable*. (Again, $\sum(r)$ depends also on x but not the criterion). Attainable boundaries are reached in a finite time with positive probability. On the contrary, the time to reach an unattainable boundary is always infinite.

Roughly speaking, $\sum(r)$ measures the time it takes to reach the boundary r starting at some interior point x . To analyze the behavior of the process if it starts at a boundary we need to introduce two more quantities:

$$M[a, r) = \lim_{b \rightarrow r} M[a, b] \tag{7.A.15}$$

and

$$N(r) = \int_x^r S[x, \xi] m(\xi) d\xi = \int_x^r M[\eta, r) s(\eta) d\eta \tag{7.A.16}$$

$M[a,r)$ measures the speed of the process near r and $N(r)$ roughly measures the time it takes to reach an interior point x starting at the boundary r .

The modern classification of boundaries is based on the boundedness or unboundedness of the four functionals $S[x,r)$, $\sum(r)$, $N(r)$ and $M[x,r)$. These functionals are not independent and there is a total of six possible combinations which correspond to different kind of boundary behavior. For our purpose we only need to consider two of them.

a) Entrance boundary

An entrance boundary cannot be reached from the interior of the state space, but it is possible, and in many applications quite natural, to consider that the process starts at the boundary. If the process starts at an entrance boundary then it quickly moves to the interior never to return to the entrance boundary.

To show that a boundary r is entrance it suffices to establish that

$$S[x,r) = \infty \tag{7.A.17}$$

while

$$N(r) < \infty \tag{7.A.18}$$

b) Natural (Feller) boundary

A natural boundary cannot be reached in finite mean time and if the process is started at such a boundary it does not reach the interior in finite mean time. Natural boundaries are omitted from the state space or are considered as separate absorbing states.

To establish that a boundary r is natural in the Feller sense one needs to show that

$$\sum(r) = \infty \tag{7.A.19}$$

$$N(r) = \infty \quad (\text{or } M[x,r]=\infty) \tag{7.A.20}$$

7.A.3 Boundaries and Stationary Measures

Only in the case of a process with two entrance boundaries does a stationary measure necessarily exist.

Lemma 7.A.3

If the two boundaries ℓ and r of the process $x(t)$ are entrance, then $x(t)$ is strongly ergodic and its stationary density is given by:

$$\psi(x) = \frac{m(x)}{\int_{\ell}^r m(\xi) d\xi} \tag{7.A.21}$$

When both boundaries are natural boundaries and, although l and r are both unattainable, neither of the possibilities

$$\text{Prob}\{ \lim_{t \rightarrow \infty} x(t) = r | x(0)=x \} = 1 \quad (7.A.22)$$

$$\text{Prob}\{ \underline{\lim}_{t \rightarrow \infty} x(t) = l, \overline{\lim}_{t \rightarrow \infty} x(t) = r | x(0)=x \} = 1 \quad (7.A.23)$$

is precluded. There may or may not exist a stationary measure approached as $t \rightarrow \infty$.

Of special interest here is the case when one of the boundaries say l , is an entrance boundary, and the other is a natural and attracting boundary. In that case, although the attracting boundary r cannot be reached in finite time, it will be reached in infinite time with probability one, i.e.,

$$\text{Prob}\{ \lim_{t \rightarrow \infty} x(t) = r | x(0)=x \} = 1 \quad (7.A.24)$$

To prove the above assertion we need the following:

Lemma 7.A.4

If $S[x,r] < \infty$ then

$$\text{Prob}\{ \lim_{t \rightarrow \infty} x(t) = r | x(0)=x \} > 0$$

Now, because l is an entrance boundary, $x(t)$ cannot converge to l with positive probability. Therefore, starting at x the process will

APPENDIX 7.B: Expected Time to Detection for the Hypothesis Testing Problem

The expected time to detection assuming $\pi_1(0) = 1/2$ is given by the following integral:

$$v_{1/2}(\delta, 1-\delta) = 2 \int_{1/2}^{1-\delta} \left(-\frac{1}{1-\delta} + \frac{1}{\xi} \right) \frac{d\xi}{K^2(1-\xi)^2} + 2\delta \int_{\delta}^{1/2} \left(-\frac{1}{\xi} + \frac{1}{\delta} \right) \frac{d\xi}{K^2(1-\xi)^2} \quad (7.B.1)$$

Because we are interested in an asymptotic evaluation of $v_{1/2}(\delta, 1-\delta)$ as $\delta \rightarrow 0$, the second term in (7.B.1) can be neglected. We thus have:

$$\begin{aligned} v_{1/2}(\delta, 1-\delta) &\sim 2 \int_{1/2}^{1-\delta} \left(-\frac{1}{1-\xi} + \frac{1}{\xi} \right) \frac{d\xi}{K^2(1-\xi)^2} = \\ &= -\frac{1}{(1-\delta)K^2} \left(\frac{1}{\delta} - 2 \right) - \frac{1}{K^2} \left[\ln \left| \frac{\delta}{1-\delta} \right| - \frac{1-\delta}{\delta} - \frac{1/2}{1/4} \right] \\ &= \frac{1}{\delta K^2} \left(1-\delta - \frac{1}{1-\delta} \right) + \frac{2}{K^2} + \frac{2}{(1-\delta)K^2} - \frac{1}{K^2} \ln \frac{\delta}{1-\delta} . \end{aligned}$$

And in the limit as $\delta \rightarrow 0$,

$$v_{1/2}(\delta, 1-\delta) = O\left(\frac{\ln(1/\delta)}{K^2}\right)$$

return to x in a finite time, unless it converges to r . At this point, the cycle starts again with a given probability of convergence to r and, if this convergence does not occur, a further finite-time excursion that will lead the process again to x . This is a sequence of Bernouilly trials which will result in $x(t)$ eventually converging to r as (7.A.24) states.

APPENDIX 7.C: Mean Exit Time from the Boundary

The diffusion under consideration has drift coefficient

$$\mu^\varepsilon(t) = \varepsilon\lambda(1-2x) + K^2 x(1-x)^2$$

and noise

$$\sigma(x) = Kx(1-x) .$$

We first compute its scale density:

$$s^\varepsilon(x) = \exp \left\{ - \int^x \frac{2\mu^\varepsilon(\tau)}{\sigma^2(\tau)} d\tau \right\} \quad (7.C.1)$$

The integral in the exponent of (7.C.1) is evaluated as follows:

$$\begin{aligned} 2 \int^x \frac{\mu^\varepsilon(\tau)}{\sigma^2(\tau)} d\tau &= \frac{2\varepsilon\lambda}{K^2} \int^x \frac{1-2\tau}{\tau^2(1-\tau)^2} d\tau + \int^x \frac{2}{\tau} d\tau \\ &= \frac{2\varepsilon\lambda}{K^2} \left[\int^x \frac{1}{\tau^2(1-\tau)^2} d\tau - 2 \int^x \frac{1}{\tau(1-\tau)^2} d\tau \right] + \ln x^2 \\ &= \frac{2\varepsilon\lambda}{K^2} \left[\frac{1}{1-x} - \frac{1}{x} - 2\ln \left| \frac{1-x}{x} \right| + 2\ln \left| \frac{1-x}{x} \right| - \frac{2x}{1-x} \right] + \ln x^2 \\ &= \frac{2\varepsilon\lambda}{K^2} \left(\frac{2x-2x^2-1}{x(1-x)} \right) + \ln x^2 \end{aligned}$$

Giving:

$$s^\varepsilon(x) = \frac{1}{x^2} \exp \left\{ - \frac{2\varepsilon\lambda}{K^2} \left(\frac{2x-2x^2-1}{x(1-x)} \right) \right\}$$

Next we establish that:

$$N^{\epsilon}(1) = \int_x^1 \int_x^{\xi} s^{\epsilon}(\eta) d\eta \cdot m^{\epsilon}(\xi) d\xi < \infty$$

where:

$$s^{\epsilon}(\eta) = \frac{1}{\eta^2} \exp \left\{ -\frac{2\epsilon\lambda}{k^2} \frac{2\eta-2\eta^2-1}{\eta(1-\eta)} \right\}$$

$$m^{\epsilon}(\xi) = \frac{1}{k^2(1-\xi)^2} \exp \left\{ \frac{2\epsilon\lambda}{k^2} \frac{2\xi-2\xi^2-1}{\xi(1-\xi)} \right\}$$

As indicated by the shape of $s^{\epsilon}(\eta)$ and $m^{\epsilon}(\xi)$ (see Fig. 7.6) the main contribution to the above integral comes from a boundary layer close to $\xi=1$ (in fact the lower limit x is arbitrary and can be taken arbitrarily close to one). The following approximations are thus warranted:

$$\frac{2\eta-2\eta^2-1}{\eta(1-\eta)} \approx -\frac{1}{1-\eta}$$

$$\frac{1}{\eta^2} \approx 1$$

Using these approximation in $N^{\epsilon}(1)$ we have:

$$N^{\epsilon}(1) \approx \int_x^1 \int_x^{\xi} \exp \left\{ \frac{2\epsilon\lambda}{k^2} \frac{1}{1-\eta} \right\} d\eta \frac{1}{k^2(1-\xi)^2} \exp \left\{ -\frac{2\lambda\epsilon}{k^2} \frac{1}{1-\xi} \right\} d\xi$$

Changing variables to

$$z = \frac{1}{1-\eta} \quad , \quad y = \frac{1}{1-\xi}$$

we get

$$N^\epsilon(1) \approx \frac{1}{K^2} \int_{\frac{1}{1-x}}^{\infty} \int_{\frac{1}{1-x}}^y \frac{1}{z^2} \exp\left\{\frac{2\lambda\epsilon}{K^2} (z-y)\right\} dz dy$$

A second change of variables in the inner integral gives:

$$N^\epsilon(1) \approx \frac{1}{K^2} \int_{\frac{1}{1-x}}^{\infty} \int_0^{y - \frac{1}{1-x}} \frac{1}{(\sigma-y)^2} \exp\left\{-\frac{2\lambda\epsilon}{K^2} \sigma\right\} d\sigma dy$$

The next step is to bound the inner integral. Pick

$0 < T < y - \frac{1}{1-x}$ and write

$$\begin{aligned} & \int_0^{y - \frac{1}{1-x}} \frac{1}{(\sigma-y)^2} \exp\left\{-\frac{2\lambda\epsilon}{K^2} \sigma\right\} d\sigma \leq \\ & \frac{1}{(T-y)^2} \int_0^T e^{-\frac{2\lambda\epsilon}{K^2} \sigma} d\sigma + e^{-\frac{2\lambda\epsilon}{K^2} T} \int_T^{y - \frac{1}{1-x}} \frac{1}{(\sigma-y)^2} d\sigma = \\ & \frac{1}{(T-y)^2} \left(1 - e^{-\frac{2\lambda\epsilon}{K^2} T}\right) \frac{K^2}{2\lambda\epsilon} + e^{-\frac{2\lambda\epsilon}{K^2} T} \left((1-x) - \frac{1}{T-y}\right) \end{aligned}$$

Choosing $T=y/2$ we get

$$N^\varepsilon(1) \leq \frac{1}{K^2} \int_{\frac{1}{1-x}}^{\infty} \left[\frac{4}{y^2} \left(1 - e^{-\frac{\lambda\varepsilon}{K^2} y} \right) \frac{K^2}{2\lambda\varepsilon} + e^{-\frac{\lambda\varepsilon}{K^2} y} \left((1-x) - \frac{2}{y} \right) \right] dy$$

and also,

$$N^\varepsilon(1) \leq \frac{2}{\lambda\varepsilon} \int_{\frac{1}{1-x}}^{\infty} \frac{1}{y^2} dy + \frac{1-x}{K^2} \int_{\frac{1}{1-x}}^{\infty} e^{-\frac{\lambda\varepsilon}{K^2} y} dy \leq M \cdot \frac{1}{\lambda\varepsilon}$$

APPENDIX 7.D: Mean Time Between False Alarms with Rare Transitions

The mean time between false alarms is given by,

$$v_{1-\delta(\epsilon)} I^{\delta(\epsilon), 1} = \int_{\delta(\epsilon)}^{1-\delta(\epsilon)} \int_{\delta(\epsilon)}^{\xi} s^{\epsilon}(\eta) d\eta m^{\epsilon}(\xi) d\xi \quad (7.D.1)$$

where $\delta(\epsilon)$ is chosen of order $O(\sqrt{\epsilon\lambda/K^2})$. We are interested here in the asymptotic behavior of $v_{1-\delta(\epsilon)} I^{\delta(\epsilon), 1}$ as $\epsilon \rightarrow 0$. Substituting $s^{\epsilon}(\eta)$ and $m^{\epsilon}(\xi)$ in (7.D.1) by their expressions we have:

$$v_{1-\delta(\epsilon)} I^{\delta(\epsilon), 1} = \int_{\delta(\epsilon)}^{1-\delta(\epsilon)} \int_{\delta(\epsilon)}^{\xi} \frac{1}{\eta^2} \exp\left\{\frac{2\epsilon\lambda}{K^2} q(\eta)\right\} d\eta \cdot \frac{\exp\left\{\frac{2\epsilon\lambda}{K^2} q(\xi)\right\}}{K^2 (1-\xi)^2} d\xi$$

where

$$q(\eta) = \frac{2\eta - 2\eta^2 - 1}{\eta(\eta-1)}$$

In the interval $\eta \in [\delta(\epsilon), 1-\delta(\epsilon)]$,

$$\exp\left\{\frac{2\epsilon\lambda}{K^2} q(\eta)\right\} = O(1) \quad \text{as } \epsilon \rightarrow 0$$

Thus

$$\begin{aligned}
 v_{1-\delta(\epsilon)} [\delta(\epsilon), 1] &= \int_{\delta(\epsilon)}^{1-\delta(\epsilon)} \int_{\delta(\epsilon)}^{\xi} \frac{1}{\eta} d\eta \cdot \frac{1}{K^2 (1-\xi)^2} d\xi = \\
 &= \int_{\delta(\epsilon)}^{1-\delta(\epsilon)} \left(\frac{1}{\delta(\epsilon)} - \frac{1}{\xi} \right) \frac{1}{K^2 (1-\xi)^2} d\xi \quad (7.D.2)
 \end{aligned}$$

and evaluating the right hand term of (7.D.2) we get:

$$\begin{aligned}
 v_{1-\delta(\epsilon)} [\delta(\epsilon), 1] &\left(\frac{1}{K^2 \delta(\epsilon)} \left[\frac{1}{\delta(\epsilon)} - \frac{1}{1-\delta(\epsilon)} \right] - \right. \\
 &\left. - \frac{1}{K^2} \left(-\ln \frac{\delta(\epsilon)}{1-\delta(\epsilon)} + \frac{1-\delta(\epsilon)}{\delta(\epsilon)} + \ln \frac{1-\delta(\epsilon)}{\delta(\epsilon)} - \frac{\delta(\epsilon)}{1-\delta(\epsilon)} \right) \right)
 \end{aligned}$$

giving the desired result

$$v_{1-\delta(\epsilon)} [\delta(\epsilon), 1] = 0 \left(\frac{1}{K^2 \delta(\epsilon)^2} \right) .$$

CHAPTER VIII: CONCLUSIONS AND SUGGESTIONS FOR
FUTURE RESEARCH

In the course of the research reported here we have developed a methodology for aggregation of linear systems with multiple time scales. Our intention has been to formalize and articulate concepts and approximations widely used in a heuristic manner in the analysis and design of large scale interconnected systems. The major contributions of this research are:

a. Conceptual Contributions

a.1) An adequate problem formulation for the analysis of time-scale phenomena and aggregation methods as a problem in singular perturbations. This problem formulation has been shown to unify and extend previous work and has proved to be of great help in providing conceptual clarification of the relationships between weak couplings, singular perturbations, time-scale separation, aggregated models and asymptotic approximations.

a.2) The notion that singular perturbation can model not only time-scale separation but also near-instabilities.

a.3) The idea that reduced-order modelling based on time-scale separation can be thought of as a problem in uniform asymptotic approximation.

a.4) The awareness that it is important to distinguish between two kinds of time scale separation: in the attenuation rate and in the frequency of oscillations.

a.5) To make evident the need to deal with stochastic discontinuity in the approximation of Markov processes with rare events.

a.6) To propose a model for the analysis of coherence phenomena in electric power networks that:

- i) links coherence to the presence of weak couplings between groups of generators;
- ii) shows that finding coherence areas is equivalent to aggregating FSMP's; and
- iii) interprets aggregated models based on coherence areas as asymptotic approximations to the system's behavior.

a.7) To point out that time-scale separation is not enough, in general, to decompose the solution of filtering (and control) problems. In addition, the structure of the observations (and control actions) available and the rates at which information about different events is collected needs to be considered.

a.8) The introduction of several qualitative performance measures for filters for FSMP's, and a criteria for judging detectors to be correct "most of the time".

b. Specific Results

b.1) A necessary and sufficient condition for a singularly perturbed LTI system to have well defined multiple time scale behavior.

- b.2) A method to compute different reduced-order models of an LTI system with well defined multiple time scale behavior.
- b.3) A uniform asymptotic approximation to $\exp\{A(\epsilon)t\}$.
- b.4) A complete characterization of stochastically discontinuous FSMP's.
- b.5) A method to produce a hierarchy of aggregated models for arbitrary singularly perturbed FSMP's.
- b.6) A method to construct an aggregated model of an electric power network based on a coherence area approximation.
- b.7) A criterion for fixing crossing levels for the detection of rare events, in such a way that the resulting detector does not have multiple false alarms and gives a correct reading "most of the time".
- b.8) A filter to detect rare transitions in singularly perturbed FSMP's that uses the aggregated model of the process to be estimated and which we show to perform near optimally "most of the time" for a simple example.

It is the author's feeling that the most important point made in this dissertation is the conceptual value of the approach taken and the clarification of the intimate relationship between weak couplings, singular perturbations, time-scale separation, aggregated models and asymptotic approximations. In establishing the results mentioned above we have used tools from the perturbation theory of linear operators in finite dimensional spaces to a far greater extent that is customary in

the System's Theory community. We hope to have brought to the attention of researchers in this field the relevance of Kato's work for the kind of problems at hand. In addition, it is hoped that future work using similar tools in an infinite dimensional setting, such as those found in [Kat 60], [Kor 78] and [Dow 80], will yield significant results extending the range of application of the ideas developed here to encompass time-varying, non-linear systems and continuous-state Markov processes. Our work on FSMP's also provides persuasive arguments that the use of analytic and asymptotic methods, as opposed to purely probabilistic ones, can be of great help in the study of Markov processes.

The following is a partial list of future research topics related to the problems and approaches discussed in this thesis.

c. Extensions and Technical Points

c.1 Reproduce the development of Chapters IV and V for discrete time systems. Study the difference equation:

$$x^\epsilon(t+1) = A(\epsilon)x^\epsilon(t) \tag{8.1}$$

We conjecture that under multiple semistability conditions, the following asymptotic expansion is valid:

$$\begin{aligned} x^\epsilon(t) &= A(\epsilon)^t x^\epsilon(0) = \\ &= A_0^t \cdot e^{A_1 \epsilon t} \cdot e^{A_2 \epsilon^2 t} \dots e^{A_m \epsilon^m t} + o(1) \end{aligned} \tag{8.2}$$

(uniformly for $t \geq 0$), where the matrices A_k are obtained using a construction similar to that used in the continuous time case.

Notice that the above expansion approximates the slow evolutions of discrete time systems by differential equations, an attractive feature for applications. See in this connection [Hop 77], [Com 76], [Jav 79] and [Del 82].

c.2) Suppose that $A_0(\epsilon)$ is uniformly stable but not MSST, i.e., suppose that the matrices A_0, A_1, \dots, A_m are all semistable except A_ℓ which has purely imaginary eigenvalues. The conjecture here is that by keeping a few extra terms in $A_\ell(\epsilon)$ we can get the following uniform asymptotic approximation:

$$\exp\{A_0(\epsilon)t\} = \exp\{A_0 t\} \dots \exp\left\{\left(\sum_{p=0}^r A_{\ell,p} \epsilon^p\right) \epsilon^\ell t\right\} \dots \exp\{A_m \epsilon^m t\} + o(1) \quad (8.3)$$

This conjecture is based on the fact that an approximation which only keeps the zeroth order terms in all the $A_k(\epsilon)$'s is not uniformly valid because the term $\exp\{A_\ell \epsilon^\ell t\}$ includes some unattenuated oscillatory modes instead of the slightly attenuated modes of the form $e^{-\epsilon^{\ell+s}} \sin \epsilon^\ell t$ that it is suppose to approximate. By keeping enough terms in the expansion of $A_\ell(\epsilon)$ this problem is corrected and the validity of the approximation extended.

c.3) Study the non-US case, i.e., systems with high amplitude transients. It is not clear if in this case a product decomposition of the form

$$\exp\{A_0(\epsilon)t\} \approx \exp\{A_0 t\} \dots \underbrace{\exp\{\tilde{A}_\ell(\epsilon) \epsilon^\ell t\}}_{\text{high amplitude transients}} \dots \exp\{\tilde{A}_m \epsilon^m t\} \quad (8.4)$$

is possible in this case and it is not clear what kind of approximation can we obtain. Some suggestions on these issues are continued in Chapter VI in connection with our discussion of aggregation of models for interconnected power systems.

c.4) Higher-order terms in the asymptotic expansion. Our results give only the dominant term in a power series expansion of $\exp\{A(\varepsilon)t\}$. To improve the approximation it may be worthwhile to compute more terms (see for example [Cam 78]). Is it true that:

$$\begin{aligned} \exp\{A_0(\varepsilon)t\} &= \prod_{k=0}^m \exp\{Q_k(\varepsilon)A_k(\varepsilon)\varepsilon^k t\} \\ &= \prod_{k=0}^m \exp\left\{\left(\sum_{p=0}^r Q_{k,p}\varepsilon^p\right) \cdot \left(\sum_{p=0}^r A_{k,p}\varepsilon^p\right) \cdot \varepsilon^k t\right\} + o(\varepsilon^r) \end{aligned} \tag{8.5}$$

c.5) Aggregated models of FSMP's with ε -dependent rates. The aggregated models we derive in Chapter VI are ε -independent. They describe the behavior of the process at a given time scale but give no clue as to what may occur at even slower time scales. It is intuitively appealing to think of ε -dependent rates *in the aggregated models* which will indicate the presence of transitions which are rare at the time-scale under consideration but that will eventually take place. This ε -corrections to the aggregated models can thus be seen as higher-order approximations of the kind mentioned in c.4). Specifically, it would be interesting to prove or disprove Delebeque's claim [Del 82] that:

$$U_k \tilde{A}_k(\varepsilon) V_k \tag{8.6}$$

is a Markov generator for $\varepsilon \in [0, \varepsilon_0]$.

d. Applications Related Tasks

d.1) In applications, it is rarely the case that the system under analysis is explicitly given as a function of a small parameter ϵ . Therefore, in order to apply the methodology proposed in this thesis a set of rules must be developed for the decomposition of a matrix A into the form $A=A_0+\epsilon B$. This is a fundamental problem in multiple time scale analysis. It must first be determined if a given matrix is amenable to time scale decomposition and if so how to effect the decomposition.

Obviously, the decomposition is not unique and some error criteria is required to judge the appropriateness of the innumerably possible combinations.

d.2) Aside from providing help in the decomposition process mentioned above, error bounds are important in their own to provide an estimate of the losses incurred when using a time-scale decomposition. An estimate of the type

$$\left\| \exp\{A_0(\epsilon)t\} - \prod_{k=1}^m \exp\{A_k \epsilon^k t\} \right\| \leq K \cdot g(\epsilon)$$

for some computable K and some function $g(\epsilon)$ will be of utmost value.

c.6) Multiple time scale analysis of singularly perturbed semigroups. Suppose now that

$$\dot{x}^\varepsilon(t) = A(\varepsilon)x^\varepsilon(t) \quad (8.7)$$

is a differential equation in a Banach space with $A(\varepsilon)$ the differential generator of a semigroup $T^\varepsilon(t)$. Under what conditions do the finite dimensional results generalize and we have a time-scale decomposition of the form:

$$T^\varepsilon(t) = T^{(0)}(t) \cdot T^{(1)}(\varepsilon t) \dots T^{(m)}(\varepsilon^m t) + o(1) . \quad (8.8)$$

for a collection of semigroups $T^{(k)}(t)$ that describe the evolution of (8.7).

c.7) Characterize stochastically discontinuous Markov processes defined on more general spaces, i.e., study the properties of positive, contraction semigroups $T(t)$ such that

$$\lim_{t \downarrow 0} T(t) = \Pi$$

Show that such a process is also uniquely characterized by its ergodic projection at zero, Π , and an aggregated version of the process.

Generalize the aggregation and disaggregation matrices U and V by

introducing the corresponding operators U, V such that $\Pi = V \cdot U$

$$U \cdot V = I.$$

c.8) Analyze the hierarchical filtering problems posed in Section 7.5. Prove or disprove the conjectures stated there.

d.3) The computational complexity of the algorithm proposed to calculate the reduced-order models A_k needs to be evaluated and Conjecture 4.4.11 should be proved or disproved. The numerical aspects of the computation of the group inverse of a matrix and of its eigenprojections also need to be studied.

REFERENCES

- [Avr 80] B. Avramovic, P. V. Kokotovic, J. R. Winkelman and J. H. Chow "Area Decomposition for Electromechanical Models of Power Systems", PAS-97, pp. 1344-1354, 1980.
- [Ben 78] A. Bensoussan, J.L. Lions and G.C. Papanicolau, "Asymptotic Analysis for Periodic Structures", North Holland, New York, 1978.
- [But 70] V. F. Butuzov, A. B. Vasil'eva and M. V. Fedoryuk, "Asymptotic Methods in the Theory of Ordinary Differential Equations", Progress in Math., Vol. 8, Plenum Press, New York, 1970.
- [Cha 81] S. M. Chan, "Small Signal Control of Multiterminal DC/AC Power Systems", PhD Thesis MIT-EECS Department, Aug. 1981.
- [Chi 76] J.S. Chipman, "Estimation and Aggregation in Econometrics: An Application of the Theory of Generalized Inverses" in Generalized Inverses and Applications, M.Z. Nasled Ed., New York, Academic Press.
- [Cam 78a] S. L. Campbell, "Singular Perturbation of Autonomous Linear Systems, II", J. Diff. Eq., Vol. 29, pp. 362-373, 1978.
- [Cam 78b] S. L. Campbell, N. J. Rose, "Singular Perturbations of Autonomous Linear Systems, III", Houston J. Math., Vol. 4, No. 4, 1978.
- [Cam 79] S. L. Campbell, N. J. Rose, "Singular Perturbation of Autonomous Linear Systems", SIAM J. Math. Anal., Vol. 10, No. 3, pp. 542-551, 1979.
- [Cam 80] S. L. Campbell, "Singularly Perturbed Differential Systems", Pitman, New York, 1980.
- [Cas 80] D. A. Castanon, M. Coderch, B. C. Levy and A. S. Willsky, "Asymptotic Analysis, Approximation and Aggregation Methods for Stochastic Hybrid Systems", Proc. JACC Conference, San Diego, California, 1980.
- [Cur 77] P.J. Curtois, "Decomposability: Queueing and Computer Systems Applications", Academic Press, New York, 1977.
- [Dav 77] E. B. Davis, "Asymptotic Analysis of Some Abstract Evolution Equations", J. Funct. Anal., Vol. 25, pp. 81-101, 1977.
- [Dav 80] E. B. Davis, "One Parameter Semigroups", Academic Press, New York, 1980.

- [Del 81] F. Delebecque, J. P. Quadrat, "Optimal Control of Markov Chains Admitting Strong and Weak Interactions", *Automatica*, Vol. 17, pp. 281-296, 1981.
- [Del 82] F. Delebecque, "A Reduction Process for Perturbed Markov Chains", submitted to *SIAM J. Appl. Math.*, 1982.
- [Doe 38] W. Doeblin, "Sur l'Equation Matricielle $A(t+s)=A(t)A(s)$ et ses Applications aux Probabilites en Chaine", *Bull. Sci. Math.*, Vol. 62, pp.21-32, 1938.
- [Doo 42] J. L. Doob, "Topics in the Theory of Markoff Chains", *Trans. Am. Math. Soc.*, Vol. 52, pp.37-64, 1942.
- [Doo 53] J. L. Doob, "Stochastic Processes" Wiley, New York, 1953, Chp. VI.
- [Dyn 65] A. A. Dynkin, "Markov Processes", Springer Verlag, Berlin 1965.
- [Eck 79] W. Eckhaus, "Asymptotic Analysis of Singular Perturbations", North Holland, Amsterdam, 1979.
- [Gai 75] V. G. Gaitsgori, A. A. Pervozvanskii, "Aggregation os States in a Markov Chain with Weak Interactions", *Kibernetika*, No. 3 pp. 91-98, 1975.
- [Gus 71] D. V. Gusak, V. S. Korolyuk, "The Asymptotic Behavior of Semi-Markov Processes with a Decomposable Set of States", *Teor. Verojatnost. i Mat. Statist.*, Vyp. 5, 1971. Translated in *Theor. Probability and Math. Statist.*, No. 5, 1975.
- [Had 71] A. H. Haddad, P. V. Kokotovic, "Note on Singular Perturbations of Linear State Regulators", *IEEE Trans. AC*, June 1971.
- [Had 77] A. H. Haddad, P. V. Kokotovic, "Stochastic Control of Linear Singularly Perturbed Systems", *IEEE Trans. Auto. Control*, AC-22, pp. 815-821, 1977.
- [Hil 57] E. Hille, R. S. Phillips, "Functional Analysis and Semigroups" *Am. Math. Soc.*, Providence, 1957.
- [Hil 77] T. Hill, "Free Energy Transduction in Biology", Wiley, New York, 1977.
- [Hop 66] F.C. Hoppensteadt, "Singular Perturbations on the Infinite Time Interval", *Trans. Amer. Math. Soc.*, Vol. 123, No. 2, pp. 521, 1966.
- [Hop 71] F.C. Hoppensteadt, "Singular Perturbations of Ordinary Differential Equations wit Small Paramenter", *Comm. Pure App. Math.*, Vol. 24, pp. 807-840, 1971.

- [Hop 75] F.C. Hoppensteadt, "Analysis of Some Problems Having Matched Asymptotic Expansion Solutions", SIAM Review Vol. 17, No. 1, 1975.
- [Kar 81] S. Karlin, H. M. Taylor, "A Second Course in Stochastic Processes", Academic Press, New York, 1981.
- [Kat 66] T. Kato, "Perturbation Theory for Linear Operators", Springer Verlag, Berlin, 1966.
- [Kei 78] J. Keilson, "Markov Chain Models - Rarity and Exponentiality", Springer Verlag, New York, pp. 132-133, 1978.
- [Kno 45] K. Knopp, "Theory of Functions", Dover, New York, 1945.
- [Kok 70] P. V. Kokotovic, R. A. Yackel, "Singular Perturbation of Linear Regulators: Basic Theorems", IEEE Trans. AC-17, 1972.
- [Kok 76] P. V. Kokotovic, R. E. O'Malley, Jr., and P. Sannuti, "Singular Perturbations and Order Reduction in Control Theory - An Overview", Automatica, Vol. 12, pp. 123-132, 1976.
- [Kok 80] P. V. Kokotovic, Ed. "Singular Perturbations and Time Scales in Modelling and Control of Dynamic Systems", Univ. of Illinois, Coordinated Science Lab., Rep. #R.901, 1980.
- [Kok 80] P.V. Kokotovic, "Subsystems, Time Scales and Multimodeling", Invited Paper, 2nd IFAC Symposium on Large Scale Systems: Theory and Applications, Toulouse, France, 1980.
- [Kor 69] V. S. Korolyuk, L.I. Polishchuk and A.A. Tomusyak, "A Limit Theorem for Semi-Markov Processes", Kibernetika, Vol. 5, No. 4, pp. 144-145, 1969.
- [Kor 70] V. S. Korolyuk, A. F. Turbin, "On the Asymptotic Behavior of the Occupation Time of a Semi-Markov Process in a Reducible Subset of States", Teor. Veroyatnost. i Mat. Statist. Vyp. 5 1970. Translated in Theor. Probability and Math. Statist., No. 2, 1974.
- [Kor 76] V. S. Korolyuk, A. F. Turbin, "Asymptotic Enlarging of Semi-Markov Processes with Arbitrary State Space", in Lecture Notes on Math. No. 550, Springer Verlag, Berlin, 1976.
- [Kor 78] V. S. Korolyuk, A. F. Turbin, "Mathematical Foundations of Phase Consolidation for Complex Systems", Naukova Dumka, Kiev, 1978, (in Russian).
- [Lag 72] P.A. Lagstrom, R.G. Casten, "Basic Concepts Underlying Singular Perturbation Techniques", SIAM Review, Vol. 14, pp.63-120, 1972.
- [Lip 78] R. S. Liptser, A. N. Shirayayev, "Statistics of Random Processes Vol II", Springer Verlag, No. 4, 1978.

- [Nay 73] A. H. Nayfeh, "Perturbation Methods", Wiley, New York, 1978.
- [Pap 75] G. C. Papanicolau, W. Kohler, "Asymptotic Analysis of Deterministic and Stochastic Equations with Rapidly Varying Components", Com. Math. Phys., Vol. 45, pp. 217- 232, 1975.
- [Per 74] A. A. Pervozvanskii, I. N. Smirnov, "Stationary State Evaluation of a Complex System with Slowly Varying Couplings", Kibernetika, No. 4, pp.45-51, 1974.
- [Phi 80] R. G. Phillips, "A Two Stages Design of Linear Feedback Controls", IEEE Trans. on Auto. Control, AC-26, pp. 1220-1222, 1981.
- [Phi 81] R. G. Phillips, P. V. Kokotovic, "A Singular Perturbation Approach to Modelling and Control of Markov Chains", IEEE Trans. on Auto. Control AC-26, pp. 1087- 1094, 1980.
- [Pod 78] R. Podmore, "Identification of Coherent Generators for Dynamic Equivalents", IEEE Trans Power App. Systems, PAS-97, pp. 1344-1354, 1978.
- [Sas 80] S. S. Sastry, C. A. Desoer, "Asymptotic Unbounded Root Loci-Formulae and Computation", to appear in IEEE Trans. AC, April 1983.
- [Sas 81] S. S. Sastry, P. Varaiya, "Coherence for Interconnected Power Systems", IEEE Trans. on AC, AC-26, pp. 218-226, 1981.
- [Sim 61] H. A. Simon, A. Ando, "Aggregation of Variables in Dynamic Systems", Econometrica, Vol. 29, No. 2, pp. 111-138, 1961.
- [Sko 65] A. V. Skorohod, "Constructive Methods of Specifying Stochastic Processes", Russian Math. Surveys, Vol. 20, pp. 70 1965.
- [Ten 77] D. Teneketzis, N. R. Sandell, Jr., "Linear Regulator Design for Stochastic Systems by Multiple Time Scale Method", IEEE Trans. on Auto. Control AC-22, pp. 615-621, 1977.
- [Ten 80] D. Teneketzis, H. Javid and B. Shridhar, "Control of Weakly-Coupled Markov Chains", Proc. CDC Conference, Albuquerque, 1980.
- [Tur 71] A. F. Turbin, "On the Asymptotic Behavior of the Sojourn Time of a Semi-Markov Process in a Reducible Subset of States: The Linear Case", Teor. Veroyatnost. i Mat. Statist., Vpr. 4, 1971. Translated in Theor. Probability and Math. Statist. No. 4, 1974.
- [Vas 63] A. B. Vasil'eva, "Asymptotic Behavior of Solutions to Certain Problems Involving Non-Linear Differential Equations Containing a small Parameter Multiplying the Highest Derivatives", Russian Math. Surveys, Vol.18, No. 3, pp. 13-84, 1963.

- [Vas 67] A. B. Vasil'eva, V. M. Volosov, "The Work of Tikhonov and His Pupils on Ordinary Differential Equations Containing a Small Parameter", Russian Math. Surveys, Vol. 22, No. 2, 1967.
- [Vas 76] A. B. Vasil'eva, "The Development of the Theory of Ordinary Differential Equations with a Small Parameter Multiplying the Highest Derivative During the Period 1966-1976", Russian Math. Surveys, Vol. 31, No. 6, pp. 109-131, 1976.
- [Vas 78] A. B. Vasil'eva, V.F. Butuzov, "Singularly Perturbed Equations in the Critical Case", Moscow State Univ., 1978, translated in Math. Res. Center, Univ. of Wisconsin, Madison, MCR-Report No. 2039.
- [Ver 81] G. C. Verghese, B. C. Levy and T. Kailath, "A Generalized State Space for Singular Systems", IEEE Trans. on Auto. Control, AC-26 pp. 811-831.
- [Vis 60] M.I. Visik, L.A. Lyusternik, "The Solution of Some Perturbation Problems for Matrices and Selfadjoint or Non-Selfadjoint Differential Equations", Russian Math. Surveys, Vol. 15, No. 3, pp. 1-73, 1960.
- [Vis 62] M. I. Visik, L. A. Lyusternik, "Regular Degeneration and Boundary Layer for Linear Differential Equations with Small Parameter", Am. Math. Soc. Transl., Series 2, Vol. 20, 1962.
- [Wal 47] W. Wald, "Sequential Analysis", Wiley, New York, 1947.
- [Was 44] "On the Asymptotic Analysis of Boundary Value Problems for Ordinary Differential Equations Containig a Parameter", J. Math. Phys., Vol. 32, pp. 173-183, 1944.
- [Was 65] W. Wasow, "Asymptotic Expansions for Ordinary Differential Equations", Wiley, New York, 1965.
- [Wil 79] D. Williams, "Diffusions, Markov Processes and Martingales Vol I; Foundations", Wiley, New York, 1979.
- [Win 80] J. R. Winkelman, J. H. Chow, P. C. Bowler, B. Avramovic and P. V. Kokotovic, "An Analysis of Interarea Dynamics of Multi-Machine Systems", presented at the IEEE PES Summer, 1980.
- [Zip 80] P.H. Zipkin, "Bounds on Aggregation of Linear Programs", Operations Research, Vol. 28, pp.403-418, 1980.